University of Cambridge

Part III of the Mathematical Tripos

Symmetries, Fields and Particles

Michaelmas 2012, Prof. N. Manton

Notes by:
William I. Jay

Diagrams by:
Ben Nachman

Last updated on June 24, 2013

Preface

I've typeset my notes from the Cambridge's Part III course Symmetries, Fields and Particles. Since my main motivation was reviewing the material for my own benefit, I've typed up everything from the lectures along with some of my own attempts to understand or rephrase the material. The (likely numerous and perhaps egregious) errors within are strictly my own. When you find such errors, please let me know at: william.jay@colorado.edu.

An enormous thank you goes to Ben Nachman for producing all of the diagrams in these notes.

Contents

Chapter 1

Introduction to Particles

The standard model incorporates the recently discovered Higgs particle. However, the standard model is still elaborate and involves many parameters. Possible next steps to makes sense of these include:

(a) Beyond-the-standard-model physics (including more particles, SUSY, or dark matter), and

(b) Simplification and unification (string theory or competitors).

Experimentally one finds many types of particles in nature, including:

- electrons
- photons
- protons / neutrons
- pions
- quarks
- gluons / glueballs
- neutrinos
- gauge particles ($W^{\pm}, Z$)
- Higgs bosons

The standard model makes detailed sense of these but is not fully understood. Experimentally, the most important properties of the observed particles are *mass* and *spin*. These are related to the *geometry* of Minkowski space. Only massless particles move at the speed of light.

The simplest theory of particles is *perturbative quantum field theory* (pQFT). In pQFT, there is one particle per field (and one particle spin state per field component). The theory is approximately linear, but this can fail when interactions between the fields are strong. Then *non-linearity* between fields becomes *crucial*. Particles associated with a field may appear as *composites* or *not at all*. Solitons are *particle-like nonlinear* field structures.

1.1 Standard Model Fields

1.1.1 Fermions: Spin 1/2 ("matter")

$$\begin{array}{cccl} \begin{pmatrix} e \\ \nu_e \end{pmatrix} & \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} & \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix} & \text{Leptons} \\ \begin{pmatrix} u \\ d \end{pmatrix} & \begin{pmatrix} c \\ s \end{pmatrix} & \begin{pmatrix} t \\ b \end{pmatrix} & \text{Quarks} \end{array}$$

Note that all fermions (except ν) have *anti-particles*. This was predicted by Dirac.

1.1.2 Bosons: Spin 0 or 1

$$\underbrace{g \text{ (gluon)}, \gamma \text{ (photon)}, W^{\pm}, Z}_{\text{spin 1}}, \quad \underbrace{H \text{ (Higgs)}}_{\text{Spin 0? Remains to be seen...}}$$

Quarks interact through gluones; leptons do not.

1.2 Observed Particles (of "long life")

- Leptons: e, ν_e, stable
- Mesons: $q\bar{q}$, for example $\pi^+ = u\bar{d}$, $\pi^- = \bar{u}d$
- Baryons:qqq, for example $p = uud$ (stable!)
- Gauge particles: γ (stable), $W^{\pm}$, Z, H, g (not seen, even in glueballs)

Remark. The strongly interacting particles are called *hadrons*: $\{mesons\} \bigcup \{baryons\} = \{hadrons\}$

1.3 Further Remarks on Particles

The pairs $\binom{e}{\nu_e}$ and $\binom{u}{d}$ lead to an $SU(2)$ structure. $SU(2)$ is a three-dimensional Lie group of 2×2 matrices which helps explain the $W^{\pm}, Z$ particles. The qqq baryons lead to a $SU(3)$ structure. $SU(3)$ is an eight-dimensional Lie group of 3×3 matrices, which explains the 8 species of gluons. The standard model has the gauge group $U(1) \times SU(2) \times SU(3)$ which extends the $U(1)$ gauge symmetry of electromagnetism with its one photon (gauge boson).

1.3.1 Mass of gauge bosons

Naively, one expects the gauge bosons in QFT to be massless. This is evaded by:

(a) *Confinement* for gluons

(b) *Higgs mechanism* for $W^{\pm}, Z$

The Higgs mechanism breaks $SU(2)$ symmetry. The $U(1) \times SU(3)$ symmetry remains unbroken.

1.3.2 The Poincare Symmetry

The *Poincare symmetry* combines translations and Lorentz transformations. The Poincare group is a ten-dimensional Lie group (think geometrically: 3 rotations, 3 boosts, and 4 translations). The Poincare symmetry explains the mass, spin, and particle-antiparticle distinction properties of particles. When Poincare symmetry is broken, particles lose definite values for mass and spin. Gravity bends spacetime, changing the Minkowski metric. Thus we expect breaking of the Poincare symmetry when gravity becomes significant.

1.3.3 Approximate Symmetries

Approximate Symmetries simplify particle classification and properties. The most important example is that $\binom{u}{d}$ have similar masses. Thus $p = uud$ and $n = udd$ have similar masses and interactions ($m_p = 938$ MeV, $m_n = 940$ MeV). This gives rise to an approximate $SU(2)$ symmetry called *isospin*. There is also another, less accurate, flavour symmetry involving the u, d, and s quarks.

1.4 Particle Models

(a) Perturbative QFT: quantize linear waves

(b) Point particles: naive quark model, non-relativistic

(c) Composites: baryons (qqq), nuclei (p, n), atoms (nuclei and e's)

(d) Exact field theory: classical localized field structures become solitons / particles after quantization

(e) String theory models of particles

We remark that multi-particle processes are hard to calculate in all models. At the LHC, $pp \longrightarrow$ hundreds of particles, mostly hadrons. Sometimes, one observes a few outgoing *jets*. These are the experimental signatures of quarks and gluons.

1.5 Forces and Processes

1.5.1 Strong Nuclear Force (quarks, gluons, $SU(3)$ gauge fields)

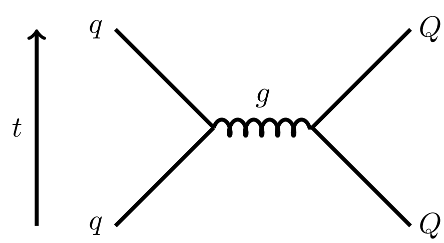


Figure 1.1: Quark Scattering, the process at the heart of hadron scattering

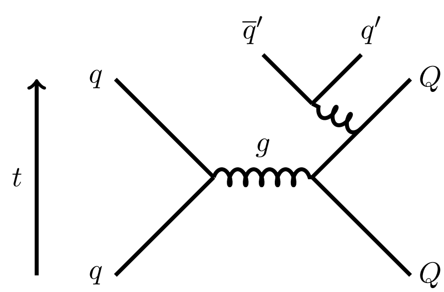


Figure 1.2: Particle Production (perhaps $pn \to pn\pi^0$)

Forces are the same for all quarks ("flavour blind"). However, the quark's masses differ: $m_\mu \sim 2-5$ MeV, $m_t \sim 175$ GeV. Note: 1 GeV $= 10^3$ MeV $\sim$ proton mass. 1 TeV $= 10^3$ GeV $\sim$ LHC energies.

Strong interactions *do NOT change quark flavour.* For each quark flavour, the net number of (quarks - antiquarks) is conserved, i.e., $N_u, N_d, N_s, N_c, N_t, N_b$ are all independently conserved in strong interactions. The total number of quarks N_q is always a multiple of three in any physical state. We write $N_q = 3B$, where B is the *baryon number*, which is evidently also conserved.

1.5.2 Electroweak Forces

Electroweak forces involve the photon γ and the vector bosons $W^\pm, Z$ and may or may not change quark flavour. However, the net number of quarks N_q remains conserved. The lepton number L is also conserved.

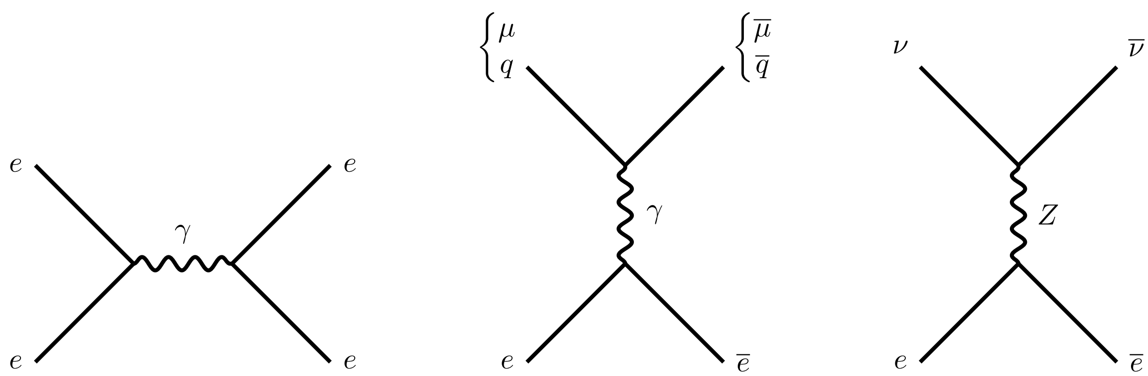


Figure 1.3: Some electroweak interactions with their Feynman diagrams

The photon only couples to electrically charged particles.

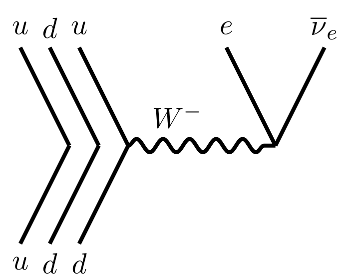


Figure 1.4: Neutron decay

Neutron decay motivated by quark decay: $n \longrightarrow pe\bar{\nu}_e$. This is a *flavour-changing electroweak process.*

Heavy meson decay involves a *transition between families* (think CKM matrix).

Heavy quark pairs ar produced in $Q\bar{Q}$ pairs in strong interactions. They separate and decay weakly, leaving short tracks. Higgs couplings determine the mass of all other particles, but do not appear

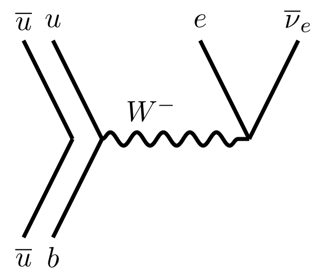


Figure 1.5: Heavy meson decay

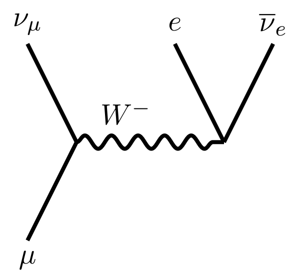


Figure 1.6: Muon decay

in Feynman scattering diagrams. The Higgs particle H shows up in diagrams like:

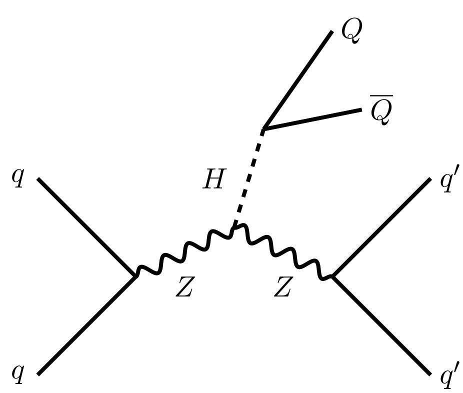


Figure 1.7: An interaction involving the Higgs

The strengths of the Higgs particle coupling are proportional to the mass of other particles, i.e., Z and Q in the case above (H couples preferentially to heavy particles).

Weak interactions are "weak" and hence slow only if the energy available is $\ll M_W, M_Z \sim 80/90$ GeV, as in neutron or muon decay. Strong interactions are "fast," occurring on time scales $\sim 10^{-24}s$ (the time for light to cross a proton).

Chapter 2

Symmetry

Definition 1. A *group* is a set $G = \{g_1 = I, g_2, g_3, \dots\}$ with

(i) a composition rule (binary operation) $g \star g'$ which we usually denote gg',

(ii) a unique identity $I \in G$ such that $Ig = gI = g$ for all $g \in G$,

(iii) an associative identity: $(gg')g'' = g(g'g'') = gg'g''$ for all $g, g', g'' \in G$, and

(iv) unique inverse: $\forall g \in G, \exists! g^{-1}$ such that $gg^{-1} = g^{-1}g = I$.

If the binary operation is commutative, we say G is *abelian*. Note that axiom (iii) is usually combined into (i) as a property of the binary operation itself.

2.1 Symmetry

Many physical and mathematical objects or physical theories possess symmetry. A symmetry is a transformation that leaves the thing unchanged. *The set of all possible symmetries forms a group*:

(i) Symmetries can be composed. We usually interpret gg' as "act with g' first and then act with g."

(ii) Doing nothing is a symmetry, the identity I.

(iii) $gg'g''$ is meaningful without brackets because of associativity

(iv) A symmetry transformation g can be reversed, which gives us the inverse g^{-1}. The inverse is itself a symmetry.

The conclusion is that group theory is the mathematical framework of symmetry. Every group is a the symmetry of something, at least itself. A natural question is then, "Why does symmetry occur in nature?" Like with most of the "big" questions, there are no easy answers. However, some partial answers are given by the following arguments:

(1) Solutions of variational problems generally exhibit a high degree of symmetry. For example, circles maximize area. As another example, Minkowski space is a stable solution of the Einstein equations (Einstein-Hilbert action) and has a high symmetry compared with a random spacetime. The symmetries of Minkowski space - known as the Poincare group - are important and discussed in some detail later in these notes.

(2) Physics often uses more mathematical variables than are really present in the physics, leading to different descriptions of the same phenomenon. Transformations between them are known as *gauge symmetries*. Gauge symmetries are *exact*. For example:

(a) Coordinate transformations

(b) Gauge transformations in electrodynamics, where the fields $\vec{B}, \vec{E}$ are physical while the potential $A^\mu = (\varphi, \vec{A}))$ is non-physical. The potential can be freely gauge transformed without altering the physics.
Remark: Gauge transformations were named by Weyl, who thought physics could not depend on "ruler." Although this ideal ultimately turned out to be wrong, as there are fundamental length scales, the name remains.

(c) Non-physical changes to a wavefunction in quantum mechanics.

(3) Approximate symmetries arise by ignoring part of the physics or by making simplifying assumptions. For example, if one ignores the differences in the small masses of the u and d quarks in hadronic physics to get the $SU(2)$ isospin symmetry.

Chapter 3

Lie Groups and Lie Algebras

Lie groups have infinitely many elements. The elements depend *continuously* of a number of (real) parameters, called the *dimension* of the group, which is usually finite. The group operations (products and inverses) depend continuously (and smoothly) on the parameters.

Definition 2. A *Lie group* G is a smooth manifold which is also a group with smooth group operations.

Definition 3. The *dimension* of G is the dimension of the underlying manifold.

The coordinates of gg' depend smoothly on the coordinates of g and g'. The inverse g^{-1} also depends smoothly on the coordinates of g.

Example:

(i) $(\mathbb{R}^n, +)$. $\mathbb{R}^n$ is a manifold of dimension n. $\vec{x}'' = \vec{x} + \vec{x}'$ is a smooth function of $\vec{x}$ and $\vec{x}'$. The inverse $\vec{x}^{-1} = -\vec{x}$ is also smooth.

(ii) $S^1 = \{\theta : 0 \leq \theta \leq 2\pi\}$ with $\theta = 0, \theta = 2\pi$ identified (to skirt the issue that some manifolds require more than one chart). Here the group operation is addition $\mod 2\pi$. Equivalently, $S^1 = \{z \in \mathbb{C} : |z| = 1\}$, with multiplication in $\mathbb{C}$ giving the group operation. The equivalence between these two portrayals can be seen via $z = \exp i\theta$. S^1 had dimension 1.

3.1 Subgroups of G

Definition 4. A *subgroup* $H \subset G$ is a subset of G closed under the group operation inherited from G. We often write $H \leq G$ to denote subgroups.

Note that a subgroup can be discrete, e.g., $\{z = 1, -1\} \leq S^1$, but discrete subgroups are *not* Lie subgroups. If H is a continuous subgroup and a *smooth submanifold* of G, then H is called a *Lie subgroup*. A Lie subgroup usually has a smaller dimension than the parent group.

3.2 Matrix Lie Groups

Lie groups of square matrices abound. These are called *linear Lie groups*, as the matrices act linearly on vectors in a vector space. The group operation in matrix Lie groups is always multiplication $M_1 \star M_1 = M_1 M_2$. Although addition is a sensible matrix operation, it is not the group operation. The identity in a matrix Lie group is always the unit matrix, and the inverse of an element M is the inverse matrix M^{-1}. Matrix multiplication is automatically associative, provided the matrix elements multiply associatively (for example, when they belong to a field). We will restrict our study to matrices over $\mathbb{R}, \mathbb{C}$ or $\mathbb{H}$. The principle example of a matrix Lie group is given by the following definition:

Definition 5. The *general Linear group* is:

$$GL(n) = \{n \times n \text{ invertible matrices}\}.$$

$GL(n, \mathbb{R})$ with entries over $\mathbb{R}$ has real dimension n^2. $GL(n, \mathbb{C})$ with entries over $\mathbb{C}$ has real dimension $2n^2$ and complex dimension n^2.

The condition of invertibility is equivalent to the condition $\det M \neq 0$. This is an "open condition," so $\dim GL(n)$ is not reduced from the dimension of the space of all $n \times n$ matrices. (Put another way, the matrices with $\det = 0$ are subset of measure zero within the space of all matrices). $GL(n)$ has a subgroup $GL^+(n, \mathbb{R}) = \{M \text{ real}, \det M > 0\}$.

3.2.1 Important Subgroups of $GL(n)$

(1) $SL(n) = M : \det M = 1$, the special linear group. Note that the group is in fact closed, since determinants of matrices multiply: $\det M_1 M_2 = \det M_1 \det M_2$.

$$\begin{aligned} &\dim SL(n, \mathbb{R}) = n^2 - 1 \text{ real dimension} \\ &\dim SL(n, \mathbb{C}) = 2n^2 - 1 \text{ real dimension, or } n^2 - 1 \text{ complex dimension} \end{aligned}$$

Note that the dimension is reduced by 1 since we've imposed one "closed" algebraic constraint.

(2) Subgroups of $GL(n, \mathbb{R})$

(i) $O(n) = \{M : M^T M = I\}$, the orthogonal group. Closure here is again easy to see, since

$$(M_1 M_2)^T M_1 M_2 = M_2^T M_1^T M_1 M_2 = I$$

$O(n)$ has inverses by construction. We note that transformations in $O(n)$ preserve length in the sense that if $\vec{v}' = M\vec{v}$ with $M \in O(n)$, then

$$\vec{v}' \cdot \vec{v}' = M\vec{v} \cdot M\vec{v} = (M\vec{v})^T M\vec{v} = \vec{v}^T M^T M\vec{v} = \vec{v} \cdot \vec{v}.$$

If $M \in O(n)$, then $\det M = \pm 1$. This follows using $M^T M = I$ and the determinant property above.

(ii) $SO(n) = \{R \in O(n) : \det R = +1\}$, the special unitary group. Physically this corresponds to the group of rotations in $\mathbb{R}^n$. If $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a frame in $\mathbb{R}^n$, then $\{R\vec{v}_1, \ldots, R\vec{v}_n\}$ is a frame *with the same orientation*. Volume elements are also preserved by $R \in SO(n)$ (in $\mathbb{R}^3$, $\vec{v}_1 \wedge \vec{v}_2 \cdot \vec{v}_3 = R\vec{v}_1 \wedge R\vec{v}_2 \cdot R\vec{v}_3$). We note that $O(n)$ additional contains orientation reversing elements, i.e., reflections, which are excluded from $SO(n)$. We note also that $\dim O(n) = \dim SO(n) = \frac{1}{2}n(n-1)$ The argument has to do with the fact that the columns of matrices in $O(n)$ are mutually orthonormal (See Example Sheet 1, Problem 3).

(3) Subgroups of $GL(n, \mathbb{C})$

(i) $U(n) = \{U \in GL(n, \mathbb{C}) : U^\dagger U = 1\}$, the unitary group. Note that $(U^\dagger)_{ij} = U^*_{ji}$. $U(n)$ preserves the norm of complex vectors, and the proof is essentially the same as for the length-preserving property of the orthogonal group. Note that $U^\dagger U = 1 \implies |\det U|^2 = 1$ (unit magnitude).

(ii) $SU(n) = \{U \in U(n) : \det U = 1\}$, the special unitary group.

$$\dim U(n) = n^2 \text{real dimension}$$
$$\dim SU(n) = n^2 - 1$$

Note that $O(n) \subset U(n)$ and $SO(n) \subset SU(n)$ are *real* subgroups.

3.2.2 A Remark on Subgroups Defined Algebraically

$GL(n)$ is obviously a smooth manifold with smooth group operations. The coordinates are the matrix elements (det $M \neq 0$ defines an open subset of $\mathbb{R}^{n^2}$ or $\mathbb{C}^{n^2}$. Subgroups defined by algebraic equations involving matrix equations (e.g. $\det U = 1$ or $M^T M = I$) are "algebraic varieties." The natural question is then, "But are they manifolds?" In general, algebraic varieties can have singularities (non-manifold points). For example, consider the Cassini ovals, defined by $(x^2 + y^2) - 2(x^2 - y^2) + 1 = b$.

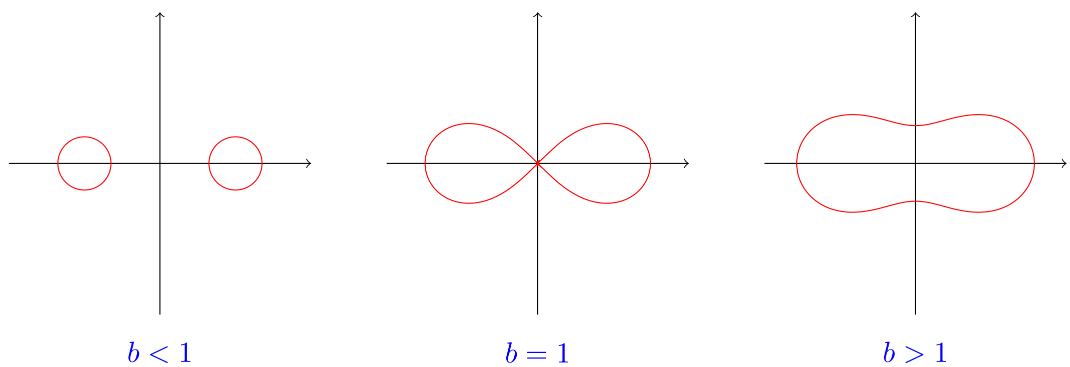


Figure 3.1: The Cassini ovals

The case $b = 1$ has a singularity. At least naively, this algebraic variety isn't a manifold. Fortunately, the *group* structure of our manifolds prevents singularities. The argument is the following:

Assume there is a singularity at $g_1 \in G$. Then there is a singularity at $g_2 \in G$, because the action of $g_2 g_1^{-1}$ by matrix multiplication is smooth. Since g_2 was arbitrary, singularities occur everywhere in G. This is a contradiction, since a variety cannot be singular everywhere. Therefore G is a smooth manifold. We conclude that algebraically defined subgroups of $GL(n)$ are Lie groups.

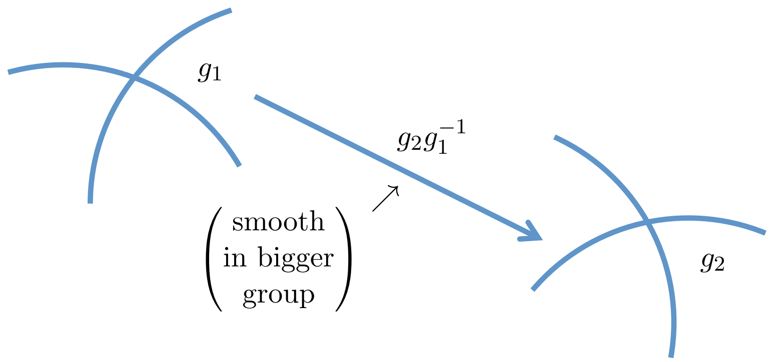
Example: $U(1) \simeq SO(2)$ These have underlying manifold S^1, the circle.

(a) $U(1) = \{\exp i\theta : 0 \leq \theta \leq 2\pi\}$, with $\theta = 0, \theta = 2\pi$ identified The product is $\exp i\theta \exp i\phi = \exp i(\theta + \phi)$.

(b) $SO(2) = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} : 0 \leq \theta \leq 2\pi \right\}$. Using trigonometric formulas, one finds $R(\theta)R(\phi) = R(\theta + \phi)$. $R(\theta)$ is an counter-clockwise rotation by the angle θ in a plane.

3.3 Lie Algebras

The Lie algebra $L(G)$ of a Lie group G is the tangent space to G at the identity $I \in G$. We study the tangent space by differentiating curves in G. $L(G)$ is a vector space of $\dim G$, with an algebraic



structure called the *Lie bracket*. The algebraic structure of $L(G)$ *almost* uniquely determines G. Group geometry thus reduces to algebraic calculations. This fact was useful to Lie and continues to be for physicists (and mathematicians, too). For two matrices X and Y, the Lie bracket is $[X, Y] = XY - YX$, i.e., the commutator. We sometimes denotes the Lie algebra of a Lie group G using the lowercase Fraktur script $\mathfrak{g}$.

3.3.1 Lie Algebra of $SO(2)$

$$g(t) = \begin{pmatrix} \cos f(t) & -\sin f(t) \\ \sin f(t) & \cos f(t) \end{pmatrix}$$

with $f(0) = 0$ is a curve in $SO(2)$ through the identity. Differentiating with respect to t and evaluating at the origin gives:

$$\left.\frac{dg}{dt}\right|_{t=0} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \left.\frac{df}{dt}\right|_{t=0}.$$

For any f, this is a multiple of $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Thus

$$\mathfrak{so}(2) = \left\{ \begin{pmatrix} 0 & -c \\ c & 0 \end{pmatrix} : c \in \mathbb{R} \right\}.$$

Note that these matrices are not *in* $SO(2)$; they're in $T_1(SO(2))$.

3.3.2 Lie Algebra of $O(n)$

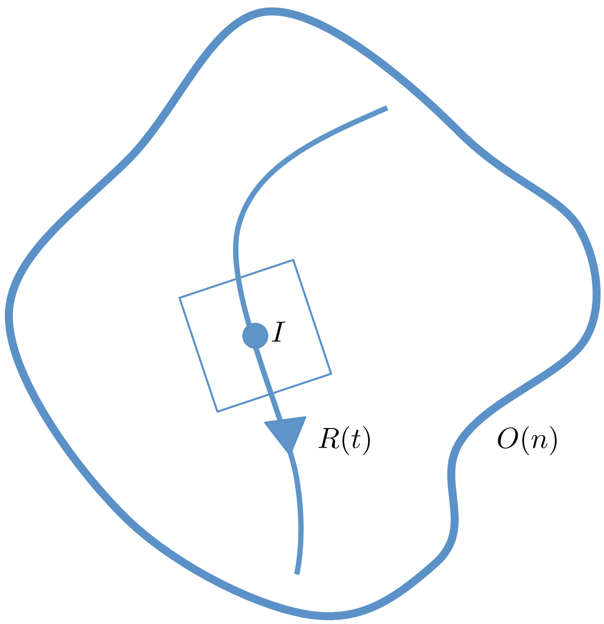
Consider a curve $R(t) \in O(n)$ with $R(t = 0) = I$. We require $R(t)^T R(t) = I$. Differentiating with respect to t, we find:

$$\frac{d}{dt}(R(t)^T R(t)) = R^T \dot{R} + \dot{R}^T R = \frac{d}{dt} I = 0, \forall t.$$

At $t = 0$, we have $R = I$, which gives the condition $\dot{R} + \dot{R}^T = 0$. In other words, $\dot{R}$ is antisymmetric. Therefore we find that

$$\begin{aligned} L(O(n)) &= \{X : X + X^T = 0\} \\ &= \{\text{vector space of real antisymmetric } n \times n \text{ matrices}\}. \\ \dim L(O(n)) &= \underbrace{1/2}_{\text{symmetry}} \ \underbrace{(n^2 - n)}_{\text{diag} = 0} = 1/2(n-1)n \end{aligned}$$

Note that $L(O(n)) = L(SO(n))$ because $O(n)$ matrices near the identity have $\det = 1$. In other words, $SO(n)$ is the part of $O(n)$ that is connected to the identity.



3.3.3 Lie algebra of $SU(n)$

Let $U(t)$ be a curve in $SU(t)$ with $U(0) = I$. Now $\det U(t) = 1, U(t)^\dagger U(t) = I$. for small t, we asume that $U(t)$ can be expanded as a power series: $U(t) = I + tZ + \dots$ with Z anti-hermitian so that $U^\dagger U = I$ to first order in t.

$$U(t) = \begin{pmatrix} 1 + tZ_{11} & tZ_{12} & \cdots \\ tZ_{21} & 1 + tZ_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

$$\begin{aligned} \implies \det U &= (1 + tZ_{11})(1 + tZ_{22}) \cdots + \mathcal{O}(t^2) \\ &= 1 + t(z_{11} + z_{22} + \dots) + \mathcal{O}(t^2) \end{aligned}$$

Thus we see that the condition $\det U = 1, \forall t \implies \operatorname{Tr} Z = 0$. Therefore

$$\begin{aligned} L(SU(n)) &= \{Z : Z + Z^\dagger = 0 \text{ and } \operatorname{Tr} Z = 0\} \\ &= \{n \times n \text{ traceless anti-hermitian matrices}\} \end{aligned}$$

3.3.4 General Structure of $L(G)$ for a matrix group G

(1) Vector Space Structure:
Suppose $X_1, X_2 \in L(G)$. Then $X_1 = \dot{g_1}(t)|_{t=0}, X_2 = \dot{g_2}(t)|_{t=0}$ for curves $g_1(t), g_2(t) \in G$ with $g_1(0) = g_2(0) = I$. Let $g(t) = g_1(\lambda t)g_2(\mu t)$ Then

$$\begin{aligned}\frac{d}{dt}(g_1(\lambda t)g_2(\mu t))|_{t=0} &= (\lambda \dot{g_1} g_2 + \mu g_1 \dot{g_2})|_{t=0} \\ &= \lambda \dot{g_1} + \mu \dot{g_2} \\ &= \lambda X_1 + \mu X_2 \\ \therefore \lambda X_1 + \mu X_2 \in L(G) &\implies L(G) \text{ is a vector space.}\end{aligned}$$

Note that since the definition of the Lie algebra is the tangent space at the identity, the computation above really just showed that our definition is consistent, since the tangent space (at any point on any manifold, not just Lie groups) is a vector space.

(2) Bracket on $L(G)$:
We'll now use more of the group structure. Let $g_1, g_2 \in G$.

$$\begin{aligned}g_1(t) &= I + tX_1 + t^2 W_1 + \dots \\ g_2(t) &= I + tX_2 + t^2 W_2 + \dots\end{aligned}$$

Computing products, we find:

$$\begin{aligned}g_1 g_2 &= I + t(X_1 + X_2) + t^2(X_1 X_2 + W_1 + W_2) + \mathcal{O}(t^3) \\ g_2 g_1 &= I + t(X_2 + X_1) + t^2(X_2 X_1 + W_2 + W_1) + \mathcal{O}(t^3)\end{aligned}$$

Define $h(t) = g_1(t)^{-1} g_2(t)^{-1} g_1(t) g_2(t)$ or equivalently $g_1(t)g_2(t) = g_2(t)g_1(t)h(t)$. We see that

$$h(t) = I + t^2[X_1, X_2] + \dots, \tag{3.1}$$

where $[X_1, X_2] = X_1 X_2 - X_2 X_1$. If we reparametrize according to $t^2 = s$, we see that $h(s)$ is a curve such that $h(0) = I$ with tangent vector $[X_1, X_2] \in L(G)$. Thus $L(G)$ is closed under the bracket operation.

Comment: Non-zero brackets are a measure of the non-commutativity of G. If G is abelian, then $L(G)$ has trivial brackets. (So in the proof above $h = I$).

Lemma 3.3.1. *If G is 1-dimensional, $L(G)$ has trivial brackets.*

Proof. $L(G) = \{cX : c \in \mathbb{R}\}$ for some matrix X. $[cX, c'X] = cc'[X, X] = 0$. □

Moreover, the only *connected* 1-dimensional Lie groups are S^1 and $\mathbb{R}$.

(3) Antisymmetry and Jacobi Identity:
The matrix bracket has the following general properties:

(a) Antisymmetry $[X, Y] = -[Y, X]$

(b) Jacobi identity: $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$
Remark: An abstract Lie algebra $\mathcal{L}$ is a vector space with the bracket $[\,,\,] : \mathcal{L} \times \mathcal{L} \to \mathcal{L}$ and satisfying antisymmetry and the Jacobi identity (cf. Humphreys, Intro. Lie Alg. and Rep. Thy.)

(4) Basis and Structure Constants:
Let $\{T_i\}$ be a basis for $L(G)$. Define structure constants by $[T_i, T_j] \equiv c_{ijk}T_k$. Antisymmetry says $c_{ijk} = -c_{jik} \Leftrightarrow c_{(ij)k} = 0$. Now computing the nested brackets gives:

$$\begin{aligned}
[[T_i, T_j], T_k] &= cijl = c_{ijl}c_{lkm}T_m \\
[[T_j, T_k], T_i] &= c_{jkl}c_{lim}T_m \\
[[T_k, T_i], T_j] &= c_{kil}c_{ljm}T_m \\
\text{Jacobi identity} \implies & c_{ijl}c_{lkm} + c_{jkl}c_{lim} + c_{kil}c_{ljm} = 0
\end{aligned}$$

3.3.5 $SU(2)$ and $SO(3)$: The Basic Non-abelian Lie Groups

We be begin by comparing the Lie algebras of $SU(2)$ and $SO(3)$, $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$

$$\mathfrak{su}(2) = \{2 \times 2 \text{ traceless, anti-hermitian matrices}\}$$

A basis is given in terms of the Pauli matrices:

$$T_a = -\frac{1}{2}i\sigma_a.$$

Here the σ_a are the Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Recall one of the properties of the Pauli matrices: $\sigma_a\sigma_b = \delta_{ab}I + i\epsilon_{abc}T_c$. Using this, one finds:

$$\begin{aligned}
[T_a, T_b] &= -\frac{1}{4}(\sigma_a\sigma_b - \sigma_b\sigma_a) \\
&= -\frac{1}{4}(i\epsilon_{abc}\sigma_c - i\epsilon_{bac}\sigma_c) \\
&= -\frac{i}{2}\epsilon_{abc}\sigma_c = \epsilon_{abc}T_c \\
\implies [T_a, T_b] &= \epsilon_{abc}T_c
\end{aligned}$$

$\mathfrak{so}(3) = \{3 \times 3 \text{ antisymmetric real matrices}\}$. A basis for $\mathfrak{so}(3)$ is given by:

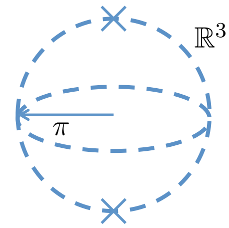
$$\tilde{T}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \tilde{T}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \tilde{T}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

In other words, $(\tilde{T}_a)_{bc} = -\epsilon_{abc}$. Then $[\tilde{T}_a, \tilde{T}_b] = \epsilon_{abc}\tilde{T}_c$, as for $\mathfrak{su}(2)$. The $\mathfrak{su}(2) \simeq \mathfrak{so}(3)$. Therefore we expect the groups $SU(2)$ and $SO(3)$ to be similar.

$SU(2)$: $U^\dagger U = I$ and $\det U = 1$ imply that U has the form (cf. Example Sheet 1, Problem 6) $U = a_0 I + i\vec{a} \cdot \sigma$ with $(a_0, \vec{a})$ real and $a_0^2 + \vec{a} \cdot \vec{a} = 1$. So the manifold on $SU(2)$ is $S^3 =$ unit sphere in $\mathbb{R}^4$.

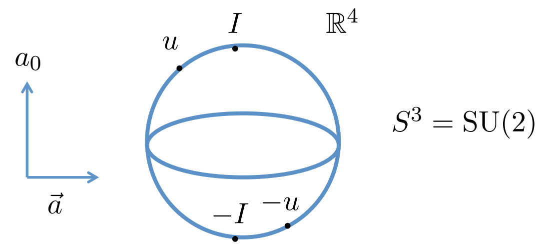
$SO(3)$: A rotation is specified by an axis of rotation $\hat{n}$ (i.e. an unit vector $\in \mathbb{R}^3$) and by an angle of rotation $\psi \in [0, \pi]$. We combine these into a 3-vector $\psi\hat{n} \in \{\text{ball of radius } \pi\} \subset \mathbb{R}^3$. Note that a rotation by π about $\hat{n}$ is equivalent to a rotation by π about $-\hat{n}$. Thus we see that opposite points on the boundary are identified. Thus the manifold does not actually have a boundary. So we see:

$$SO(3) = \{\text{ball in } \mathbb{R}^3 \text{ of radius } \pi \text{ with opposite points on boundary identified}\}$$

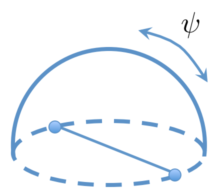


3.3.6 The Isomorphism $SO(3) \simeq SU(2)/\mathbb{Z}_2$

$SU(2)$ has a center $Z(SU(2)) = \mathbb{Z}_2 = \{I_2, -I_2\}$. If $u = a_0 I + i\vec{a}\cdot\sigma$, then $(-I)u = -a_0 I - i\vec{a}\cdot\sigma$. Thus $SU(2)/\mathbb{Z}_2 = S^3$ with antipodal points identified.



$$\begin{aligned}\therefore\ SU(2)/\mathbb{Z}_2 &= \{\text{upper half of } S^3\ (a_0 \geq 0) \text{ with opposite points of } S^3\text{-equator identified}\}\\ &= \{\text{curved version of } S0(3)\}.\end{aligned}$$

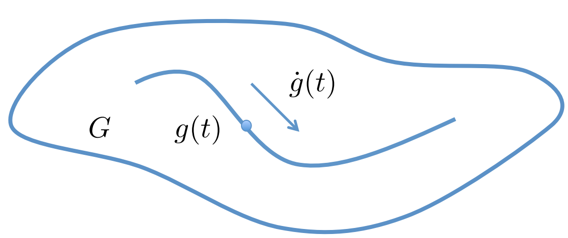


Note that the "curvature" is immaterial (topologically it is a refinement?). There is an explicit correspondence $U \in SU(2) \mapsto R(U) \in SO(3)$ with $R(-U) = R(U)$, where $U = \cos\frac{\alpha}{2} + i\sin\frac{\alpha}{2}\hat{n}\cdot\vec{\sigma} \mapsto R(U) =$ rotation by α about $\hat{n}$.

3.4 Lie Groups

3.4.1 Curves in Lie Groups

Let G be a matrix Lie group and $g(t) \in G$ a curve. $\frac{dg}{dt} \equiv \dot{g}$ is the tangent (matrix) at $g(t)$ and $g(t+\epsilon) = g(t) + \epsilon\dot{g}(t) + \mathcal{O}(\epsilon^2)$, where ϵ is infinitesimal. We can also write $g(t+\epsilon)$ as a product in G: $g(t+\epsilon) = g(t)h(\epsilon)$, where $h(\epsilon) = I + \epsilon X + \mathcal{O}(\epsilon^2)$ for some $X(t) \in L(G)$ is the group element that generates the translation $x \mapsto x + \epsilon$.



Then $I + \epsilon X = h(\epsilon) = g(t)^{-1}(g(t) + \epsilon \dot{g}(t)) = I + \epsilon g(t)^{-1}\dot{g}(t)$ to $\mathcal{O}(\epsilon)$. So $X(t) = g(t)^{-1}\dot{g}(t)$. Thus:

$$g(t)^{-1}\dot{g}(t) \in L(g), \forall t. \tag{3.2}$$

Similarly, we also have $\dot{g}(t)g(t)^{-1} \in L(G)$ (although in general $\neq g^{-1}\dot{g}$). These statements allow us "to go from the group to the algebra." We now consider the converse. Suppose $X(t) \in L(G)$ is a given curve in the Lie algebra. We can write down the equation:

$$\begin{aligned}
&g(t)^{-1}\dot{g}(t) = X(t), \\
&g(0) = I \text{ (could be the more general element } g_0).
\end{aligned}$$

This equation with an initial condition has a unique solution in G. The equation makes intuitive sense because "velocity" is always tangent to G. As a special case, consider $X(t) = X = \text{const}$. Then $\dot{g} = gX$ with $g(0) = I$. This equation has the solution $g(t) = \exp tX$.

Proof.

$$\begin{aligned}
\exp tX &= \sum_{n=0}^{\infty} \frac{1}{n!}(tX)^n \\
\frac{d}{dt}(\exp tX) &= X + tX^2 + \frac{1}{2}t^2X^3 \\
&= \exp(tX)X
\end{aligned}$$

Thus the equation is satisfied and $g(0) = I$. Note also that $g(t) = \exp(tX)$ commutes with X, so also solve the equation $\dot{g} = Xg$. □

Claim: The curve $\{g_x(t) = \exp(tX) : \infty < t < \infty\}$ is an abelian subgroup of G, generated by X.

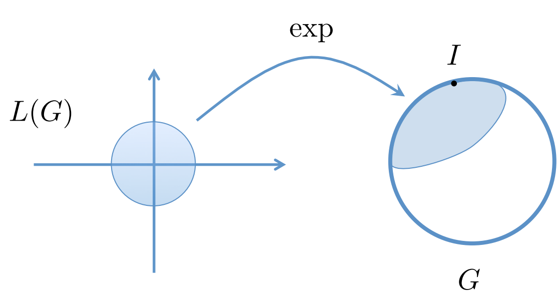
Proof.

$$\begin{aligned}
g(0) &= I, \\
g(s)g(t) &= g(t+s) = g(s+t) \text{ since X commutes with itself,} \\
g(t)^{-1} &= g(-t)
\end{aligned}$$

□

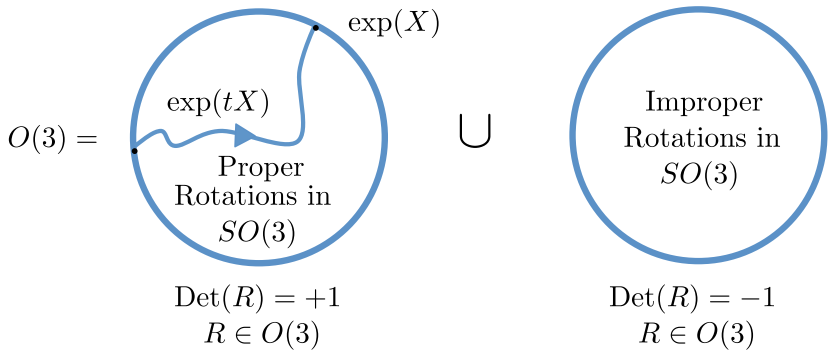
(So this tells us that $g_x(\,)$ is isomorphic either to $(\mathbb{R}, +)$ if $g_x(t) = I$ only for $t = 0$ or to S^1 if $g_x(t_0) = I$ for some $t_0 \neq 0$ and not for all t.

We can gain a general insight from the above considerations. Set $t = 1$, and consider all $X \in L(G)$. We have then found a map $L(G) \to G, X \mapsto \exp X$. This map is locally bijective (proof omitted), as all elements $g \in G$ close to I can be expressed uniquely as $\exp X$ for some small X.



Note that this map is *not* globally simple and in most cases not even one-to-one. For example, $\exp i\theta \in \{z \in \mathbb{C} : |z| = 1\}$ is massively not one-to-one, e.g., $\exp i2\pi n = 1, \forall n \in \mathbb{Z}$.
The image of exp is not the whole group, but rather the *component connected to I* in G.

Example: $O(3)$



So $\exp \mathfrak{so}(3) = \exp \mathfrak{o}(3) = SO(3)$. Evidently improper rotations cannot be expressed as $\exp X$ with X antisymmetric and real.

3.4.2 The Baker-Campbell-Hausdorff Formula

If group elements are expressed as $\exp X, X \in L(G)$, can we calculate products? Remarkably, there is a universal formula:

$$\exp X \exp Y = \exp Z, \text{ where } Z = X + Y + \frac{1}{2}[X, Y] + \text{higher brackets.}$$

We deduce that the Lie algebra, with it's bracket structure, determines the group structure of G near I.

Remark: Suppose $g(x_1, x_2, \ldots, x_k$ is a G-valued function on $\mathbb{R}^k$. Then $\frac{\partial}{\partial x^i} g \equiv \partial_i g$ is in the tangent space to G at g, so $(\partial_i g)g^{-1} \in L(G)$ and $g^{-1}(\partial_i g) \in L(G)$. These formulas appear, for example, in gauge theory.

Chapter 4

Lie Group Actions: Orbits

A Lie group G can act in many ways on other objects.

Definition 6. An *action* of G on a manifold $\mathcal{M}$ is a set of maps $g : \mathcal{M} \to \mathcal{M}$ for all $g \in G$, consistent with the group properties $g_1(g_2(m)) = (g_1 g_2)(m)$ for all $g_1, g_2 \in G$ and for all $m \in \mathcal{M}$.

Note: This is equivalent to a map $G \times \mathcal{M} \to \mathcal{M}$, assumed here to be smooth in both arguments.

Definition 7. The *orbit* of a point $m \in \mathcal{M}$ is the set $G(m) = \{g(m) : g \in G\}$

Proposition 1. *If $m' \in G(m), G(m) = G(m')$*

Proof. $m' = g(m) \implies G(m') = G(g(m)) = (Gg)(m) = G(m)$ $\square$

Theorem 4.0.1. *$\mathcal{M}$ is a disjoint union of orbits*

Proof. Omitted. $\square$

Example. $SO(n)$ acts of $\mathbb{R}^n$. The orbits are the spheres S^{n-1}, labelled by the radius r. The origin $\vec{0} \in \mathbb{R}^n$ is a special points, as its orbit is a single point.

4.1 Examples of Group Actions

Definition 8. The *left action* of G on G is $g : G \to G$ with $g(g') = gg'$

Definition 9. The *right action* of G on G is $g : G \to G$ with $g(g') = g'g^{-1}$

Note that the right action uses the inverse of g in the definition of its function; this is necessary so that the action satisfies the group law. Each of these actions is *transitive* on G.

Definition 10. An action $G \times \mathcal{M} \to \mathcal{M}$ is said to be *transitive* if $\mathcal{M}$ consists of one orbit.

Transitivity of the left and right actions. Let $g', g'' \in G$. g' and g'' are in the same left orbit, since $g(g') = g''$ when $g = g''g'^{-1}$. A similar argument applies for right orbits. $\square$

Definition 11. *Conjugation* is the action $g : G \to G$ defined by $g(g') = gg'g^{-1}, \forall g, g' \in G$

The orbit structure under conjugation is more interesting. In particular, one can show (cf. Example Sheet 2, Problem 4) that "conjugation preserves eigenvalues." The idea is that conjugation amounts to a change of basis in a matrix Lie group. Note also that one orbit is the identity alone, since $g(I) = gIg^{-1} = I, \forall g$.

Definition 12. The *combined action* of $G \times G$ on G is defined by $(g_1, g_2)(g') = g_1 g g_2^{-1}$.

Note that $G_1 \times G_2 = \{(g_1, g_2) : g_1 \in G_1, g_2 \in G_2\}$ with the product given by $(g_1, g_2) \cdot (g_1', g_2') = (g_1 g_1', g_2 g_2')$ and identity (I_1, I_2).

Example. The action of $SU(2)$ on $SU(2)$.

Recall that we can express $g \in SU(2)$ as $g = a_0 I + i\vec{a} \cdot \vec{\sigma}$ with the constraint $a_0^2 + \vec{a} \cdot \vec{a} = 1$ (One way to see this parametrization is to think of $SU(2)$ as the subgroup of the quaternions $\mathbb{H}$ with unit length). This parametrization leads us to understand $SU(2)$ to be S^3 as a manifold. If $g = a_0 I + i\vec{a} \cdot \vec{\sigma}$ and $g' = b_0 I + i\vec{b} \cdot \vec{\sigma}$, we can look at the left action given by:

$$\begin{aligned} g(g') \equiv gg' &= (a_0 b_0 - \vec{a} \cdot \vec{b})I + i(a_0 \vec{b} + \vec{a} b_0 - \vec{a} \times \vec{b}) \cdot \vec{\sigma} \\ &\equiv c_0 I + i\vec{c} \cdot \vec{\sigma} \end{aligned}$$

(cf. Example Sheet 1, Problem 6 for details.) We see that $(c_0, \vec{c})$ depends linearly on $(b_0, \vec{b})$ and $c_0^2 + \vec{c} \cdot \vec{c} = b_0^2 + \vec{b} \cdot \vec{b} = a_0^2 + \vec{a} \cdot \vec{a} = 1$. Thus we see that the left action of g defines an element of $O(4)$.

The identity I acts trivially, and $SU(2)$ is connected, so the left action of g must be a proper rotation (a proper rotation is an element of $SO(4)$ and has det $= +1$). We deduce that $SU(2)_L \leq SO(4)$, where the notation $\leq$ denotes a subgroup. Similarly, the right action gives a different subgroup $SU(2)_R \leq SO(4)$. In fact, the combined action of $SU(2)_L \times SU(2)_R$ gives every element of $SO(4)$.

Theorem 4.1.1. $SO(4) \simeq SU(2)_L \times SU(2)_R / \mathbb{Z}_2$

Proof. Omitted (cf. Example Sheet) □

(The subgroup $\mathbb{Z}_2$ consists of (I, I) and $(-I, -I)$, since (g_1, g_2) and $(-g_1, -g_2)$ act identically as $SO(4)$ transformations. Because of the group structure given in the theorem, it follows that the Lie algebra is given by $\mathfrak{so}(4) = \mathfrak{su}(2)_L \oplus \mathfrak{su}(2)_R$.)

4.2 The General Nature of an Orbit of G

Definition 13. Let G act on $\mathcal{M}$ transitively (single orbit). Let $m \in \mathcal{M}$. The *isotropy subgroup* (or *stabilizer*) at m is the subset H of G that leaves m fixed:

$$H = \{h \in G : h(m) = m\}$$

We often write $H(m) = m$. H is a subgroup of G, because if $h_1(m) = m, h_2(m) = m$ for $h_1, h_2 \in H$ and $m \in \mathcal{M}$, then $(h_1 h_2)(m) = h_1(h_2(m)) = h_1(m) = m$. The inverses and identity are trivial.

Using m as a base point for $\mathcal{M}$, we can identify another point m' with a *coset* of H. If $m' = g(m)$, then $m' = gH(m)$. The element m' is identified not just with one element g that sends m to m' but with the whole (left) coset gH. Thus we have

$$\mathcal{M} \simeq \{\text{space of all cosets of } H \text{ in } G\} \equiv G/H$$

However, there is nothing special about the base point m. The isotropy group at m' is $H' = gHg^{-1}$, and this is structurally the same as H. We can check this:

$$H'(m') = gHg^{-1}(m') = gHg^{-1}g(m) = gH(m) = g(m) = m'$$

Thus we regard G/H and G/H' as the same. This motivates the following definition:

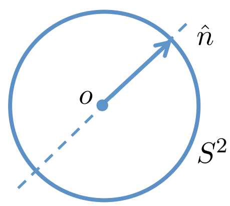
Definition 14. Let G be a group acting continuously and transitively on the manifold $\mathcal{M}$. $\mathcal{M}$ is said to be a *homogeneous space*.

The critical part of the definition above is the transitivity: since there's is just one orbit, all points on $\mathcal{M}$ are "similar."

Proposition 2. *If H is a Lie subgroup of G and $\mathcal{M}$ is as above, then* $\dim \mathcal{M} = \dim G - \dim H$.

We can check this statement near m. The tangent space to $\mathcal{M}$ is $L(G)/L(H)$ (as vector spaces), as H acts trivially. If we find a vector space decomposition $L(G) = L(H) \oplus V_m$, V_m is a vector space complement to $L(H)$ (could be orthogonal complement). V_m can be identified with the tangent space to $\mathcal{M}$ at m.

Example. $SO(3)$ acts transitively on S^2, the unit sphere, since any point can be rotated into any other point. The isotropy group at $\hat{n}$ is the $SO(2)$ subgroup of $SO(3)$ of rotations about the axis through $\vec{0}$ and $\hat{n}$. Thus $S^2 = SO(3)/SO(2)$. Note that we usually choose $\hat{n}$ so that the $SO(2)$ rotations are about the x_3-axis.



Chapter 5

Representations of Lie Groups

Definition 15. A *representation* $D(G)$ of G is a linear action of G on a vector space V. Let $\dim V = N$. Then N is called the *dimension* of D and $D(g) \in GL(N), \forall g \in G$.

Note that in order to get explicit matrices me must choose a basis for V. Linearity says that

$$D(g)(\alpha v_1 + \beta v_2) = \alpha D(g)v_1 + \beta D(g)v_2$$

That the representation is a group action means: $D(g_1 g_2) = D(g_1)D(g_2)$. We see that identity and inverses are particularly well-behaved: $D(I) = I_N$ and $D(g^{-1}) = D(g)^{-1}$.

Definition 16. A representation is *faithful* if $D(g) = I_N$ ONLY for $g = I$.

In a faithful representation distinct group elements are represented by distinct matrices. Note that a slightly more sophisticated definition would say that the homomorphism $\phi : G \to GL(N)$ is injective, i.e., that $\ker(\phi)$ is trivial.

Example. Representations of the additive group $\mathbb{R}$. We require $D(\alpha + \beta) = D(\alpha)D(\beta)$.

(a) $D(\alpha) = \exp(k\alpha), k \in \mathbf{R}$, faithful if $k \neq 0$

(b) $D(\alpha) = \exp(ik\alpha), k \in \mathbf{R}$, not faithful

(c) $D(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$, $N = 2$, not faithful

(d) Let $V = \{\text{space of functions of } x\}$. $(D(\alpha)f)(x) = f(x + a)$, infinite dimensional

5.1 Types of Representations

Definition 17. Let G be a matrix Lie group of dimension N. Then the representation that takes $D(g) = g$ is called the *fundamental representation* of the group.

Note: The notion of the fundamental representation may also be applied to more general Lie groups, but one doesn't frequently encounter it in physics.

Definition 18. The representation of G that takes $D(g) = I_N, \forall g \in G$ (and for any N) is called the *trivial representation*.

Definition 19. Let G be a Lie group. Let $V = L(g)$, the Lie algebra of G. The *adjoint representation* of G, denoted Ad, is the natural representation of G on its own Lie algebra:

$$D(g) \equiv Ad_g(X) \equiv gXg^{-1}, g \in G, X \in L(G)$$

Note that the adjoint representation Ad is the linearized version of the action of G on itself by conjugation. We do a couple of checks to make sure that we have a well-defined representation:

- Closure: $gXg^{-1} \in L(G)$
  There exists some curve $g(t) = I + tX + \ldots$ in G with tangent X at $t = 0$. Then $\tilde{g}(t) = gg(t)g^{-1}$ is another curve in G and $\tilde{g}(t) = I + tgXg^{-1} + \ldots$ with tangent gXg^{-1} at $t = 0$, thus $gXg^{-1} \in L(G)$.

- Ad is a representation:

$$\begin{aligned}(\mathrm{Ad}\, g_1 g_2)X &= g_1 g_2 X (g_1 g_2)^{-1} \\ &= g_1 g_2 X g_2^{-1} g_1^{-1} \\ &= (\mathrm{Ad}_{g_1})(\mathrm{Ad}_{g_2})X\end{aligned}$$

Note that we are thinking of $L(G)$ as a real vector space, so $\mathrm{Ad}\, G \in GL(\dim G, \mathbb{R})$. In fact, for $U(n)$ and $O(N)$, $\mathrm{Ad}\, g \in SO(\dim G)$.

Definition 20. An N-dimensional representation D of G is said to be *unitary* if $D(g) \in U(N), \forall g \in G$. If D is also real, then $D(g) \in O(N)$, and the representation is said to be *orthogonal*.

Remark: Unitary representation are important in quantum mechanics and its various generalizations because a symmetry group should preserve the *norm* of all wave functions.

Definition 21. Let D be a representation of G acting on the vector space V. Let A be an invertible transformation on V. Then we say that $\tilde{D} = AD(g)A^{-1}$ is an *equivalent representation* of G.

Note that equivalent representations are related by a change of basis of the vector space V.

Definition 22. Let D be a representation of G acting on V. D is *reducible* if there exists a proper, invariant subspace $W \subset V$, i.e., there exists a subspace W such that $D(G)W \subseteq W$. If no such subspace exists, then D is an *irreducible representation*, which we will sometimes call an "irrep."

Definition 23. A representation D is *totally reducible* if it can be decomposed into irreducible pieces, i.e., if there exists a (possibly infinite) direct sum decomposition $V = W_1 \oplus W_2 \oplus \ldots W_k$ such that $D(G)W_i \subseteq W_i$ and D restricted to W_i is an irrep.

Note that we in fact have $D(G)W = W$, since $I \in G$. Also, for a fixed $g \in G$, $D(g)W = W$, because $D(g)$ is invertible. In matrix language, for a totally reducible representation there exists a basis for V such that $D(g)$ is block diagonal, taking the form:

$$D(g) = \begin{pmatrix} D_1(g) & 0 & \cdots & 0 \\ 0 & D_2(g) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & D_k(g) \end{pmatrix},$$

where the D_i are each irreps.

Theorem 5.1.1. *A finite dimensional unitary representation is totally reducible*

Proof (sketch). For each invariant subspace W, the orthogonal complement $W_\perp$ is also invariant, and so $V = W \oplus W_\perp$. Now reduce W and $W_\perp$ until the process ends. Not that the process must end, since the theorem assumes the representation to be finite dimensional.
(cf. Example Sheet 2, Problem 9 for more details) □

These notions are important, because vectors within W_i are actually related by G. If G is a symmetry, physical states (particles!) in an *irreducible* W have similar properties.

Chapter 6

Representations of Lie Algebras

By restricting a representation D of G to elements close to the identity I, we obtain the notion of a representation of the Lie algebra $L(G)$.

Definition 24. A *representation* d *of* $L(G)$ acting on a vector space V is a linear action $v \mapsto d(X)v$, with $X \in L(G)$ and $v \in V$, satisfying $d([X,Y]) = d(X)d(Y) - d(Y)d(X) = [d(X), d(Y)]$. As with the case of groups, the *dimension* of the representation is $\dim d = \dim V = N$, where N is the dimension of the vector space V.

Example: For a matrix Lie algebra, the fundamental representation is $d(X) = X$. There also exists the trivial representation in which $d(X) = 0, \forall X \in L(G)$.

A representation d of $L(G)$ is called (anti-)hermitian if $d(X)$ is (anti-)hermitian for all $X \in L(G)$.

6.1 Representation of $L(G)$ from a Representation of G

Let $g(t) = I + tX + \cdots \in G$. Write $D(g(t)) = I_N + td(X) + \ldots$, which defines $d(X)$. Then d is the representation of $L(G)$ associated to D. We check now that the Lie bracket is preserved:

$$D(g_1^{-1}g_2^{-1}g_1g_2) = D(g_1)^{-1}D(g_2)^{-1}D(g_1)D(g_2),$$

with

$$\begin{aligned} g_1(t) &= I + tX_1 + t^2W_1 + \ldots \\ g_2(t) &= I + tX_2 + t^2W_2 + \ldots \\ g_1^{-1}(t) &= I - tX_1 - t^2W_1 + \ldots \\ g_2^{-1}(t) &= I - tX_2 - t^2W_2 + \ldots \end{aligned}$$

After a brief computation very similar to the one leading to (3.1) we see that:

$$\begin{aligned} LHS &= D(I + t[X_1, X_2] + \ldots) = I_N + t^2 d([X_1, X_2]) \\ RHS &= I_N + t^2[d(X_1), d(X_2)] + \ldots, \end{aligned}$$

and so we see the bracket is preserved.

Note: If D is a unitary representation of G, then $d(X)$ is antihermitian for all $X \in L(G)$. Why is this true? If $D(X)$ is unitary, then $D(X) = I + d(X)$ and $D(X)^\dagger = I + td(X)^\dagger$, with $d(X) + d(X)^\dagger = 0$.

thus $d(X)$ is antihermitian.

Example: The representation ad of $L(G)$ associated to Ad of G (Both of these representations act on $L(G)$, though one is a group representation and the other is an algebra representation). Recall that the adjoint representation $\mathrm{Ad}_g Y = gYg^{-1}, Y \in L(G)$. Set $g(t) = I + tX$ and $g^{-1}(t) = I - tX$. Then

$$\begin{aligned}
\mathrm{Ad}_g Y &= (I + tX)Y(1 - tX) + \mathcal{O}(t^2) \\
&= IYI - IYtX + tXYI + \mathcal{O}(t^2) \\
&= Y + t[X, Y] + \mathcal{O}(t^2) \\
&\equiv (I + t(\mathrm{ad}_X))Y.
\end{aligned}$$

Thus we see that $\mathrm{ad}_X Y = [X, Y]$. This is the adjoint representation of $L(G)$ acting on itself. We can check easily (about three lines of computations) that $\mathrm{ad}[X, Y] = [\mathrm{ad}\, X, \mathrm{ad}\, Y]$ using the Jacobi identity.

6.2 Representation of G from a Representation of $L(G)$

Given $g \in G$, express g as $\exp X$ for $X \in L(G)$, then use the formula: $D(\exp X) = \exp(d(X))$. At least locally (i.e., in a neighborhood of the identity), this defines D using the exponential function to move from the Lie algebra up to the group. How do we check this result? We need to shows that D "commutes appropriately" with exp:

$$\begin{aligned}
&D(\exp X)D(\exp Y) = D(\exp X \exp Y) \\
&D(\alpha \exp X + \beta \exp Y) = \alpha D(\exp X) + \beta D(\exp Y)
\end{aligned}$$

This computation is done in Example Sheet 2, Problem 11. In practice (in physics), we representations of $L(G)$ corresponding to representations of G:

(a) Always true: Rep. of $G \longrightarrow$ Rep. of $L(G)$

(b) Mostly true: Rep. of $L(G) \longrightarrow$ Rep. of G (works locally, but can encounter problems globally)

6.3 $\mathfrak{su}(2)$: The Mathematics of Angular Momentum Theory

From our knowledge of angular momentum theory, we know that $\mathfrak{su}(2)$ has the standard basis:

$$\{T_a = -\frac{1}{2} i\sigma_a : a = 1, 2, 3\}.$$

It is convenient for our current purposes to construct a new basis using (non-real) linear combinations:

$$\begin{aligned}
h &= \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} = \frac{1}{2}\sigma_3 = iT_3 \\
e_+ &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}(\sigma_1 + i\sigma_2) = iT_1 - T_2 \\
e_- &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \frac{1}{2}(\sigma_1 - i\sigma_2) = iT_1 + T_2
\end{aligned}$$

Note: This notation with h and $e_\pm$ appears in the text by Humphreys
Remark: h, e_+, e_- can be thought of as analogous to the familiar operators J_3, J_+, J_- in angular momentum theory. We have the brackets:

$$\begin{aligned}[h, e_+] &= e_+\\ [h, e_-] &= -e_-\\ [e_+, e_-] &= 2h.\end{aligned}$$

One sees that we can express these commutator relations in terms of the adjoint representation ad of $\mathfrak{su}(2)$:

$$\begin{aligned}(\operatorname{ad} h)e_+ &= e_+\\ (\operatorname{ad} h)e_- &= -e_-\\ (\operatorname{ad} h)h &= 0.\end{aligned}$$

What's more, we see that we have diagonalized the operator ad_h, and e_+, e_- are eigenvectors. What's going on here in terms of the broader theory? The maximal commuting (i.e., abelian) subalgebra in $\mathfrak{su}(2)$ is generated by h. From the text by Costa and Folgi, we have the following definition:

Definition 25. Let L be a semisimple Lie algebra. The *Cartan subalgebra* is the subalgebra H of L such that:

(a) H is the maximal Abelian subalgebra of L

(b) The adjoint representation ad_h $(h \in H)$ of H is completely reducible.

Definition 26. Let H be the Cartan subalgebra of L. The nonzero eigenvalues of $\operatorname{ad}_h, h \in H$ are called *roots*.

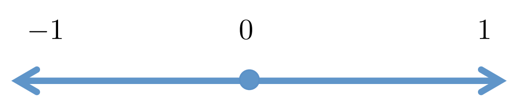


Figure 6.1: The root diagram for $L(SU(2))$

6.3.1 Irreducible Representations of $\mathfrak{su}(2)$

The irreducible representations of $\mathfrak{su}(2)$ are denoted $d^{(j)}$, with the label j (spin) as follows:

j	0	1/2	1	3/2	...
dim	1	2	3	4	...

The representation $d^{(j)}$ acts on $V^{(j)}$, a vector space of dim $= 2j+1$. We introduce a basis for $V^{(j)}$: $\{|j, m\rangle\}$, with $m = j, j-1, \ldots, -j+1, -j$ ($2j+1$ total numbers). Then:

$$\begin{aligned}d^{(j)}(h)\,|j, m\rangle &= m\,|j, m\rangle\\ d^{(j)}(e_-)\,|j, m\rangle &= \sqrt{(j-m+1)(j+m)}\,|j, m-1\rangle\\ d^{(j)}(e_+)\,|j, m\rangle &= \sqrt{(j-m+1)(j+m)}\,|j, m+1\rangle\end{aligned}$$

One can show that $[d^{(j)}(e_+), d^{(j)}(e_-)] = 2d^{(j)}(h)$ and so on for the other commutators. This irreducible representation is antihermitian, since $d^{(j)}(h)^\dagger = d^{(j)}(h)$ and $d^{(j)}(e_+)^\dagger = d^{(j)}(e_-)$, which tells us that $d^{(j)}(T_a)$ is an antihermitian matrix. (Note: there's something slightly weird going on here, since the basis $\{h, e_+, e_-\}$ lives in the complexification $\mathfrak{su}(2)^{\mathbb{C}} = \mathfrak{su}(2) \otimes_{\mathbb{R}} \mathbb{C}$). The eigenvalues of $d^{(j)}(h)$ are called *weights*. In this example, the weights are the given by the labels m.

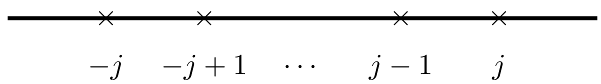


Figure 6.2: The weight diagram of $d^{(j)}(h)$

For some specific representations, we have: $d^{(0)}$ is the trivial representation, $d^{(\frac{1}{2})}$ is the fundamental representation, and $d^{(1)}$ is the adjoint representation. We note that $d^{(j)}$ exponentiates to an irreducible representation $D^{(j)}$ of $SU(2)$.

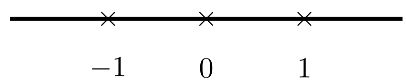


Figure 6.3: The weight diagram of $d^{(1)}$, which is the same as the root diagram including zero

For the sake of clarity we collect the following points regarding roots and weights:
Weights vs. Roots: Given a semi-simple Lie algebra L of dimension n, one can find a standard basis $\{h_1, h_2, \ldots, h_k; e_1, e_2, \ldots, e_{n-k}\}$, where the h_i span the Cartan subalgebra. One then considers the various irreducible representations $d(L)$ of the Lie algebra and assumes that the $d(h_i)$ are simultaneously diagonalized (this always works). Let $d(L)$ act on the N-dimensional vector space V, spanned by the basis $\{\psi_1, \psi_2, \ldots, \psi_N\}$. Consider one of the basis vectors ψ, which by construction satisfies

$$h_i \psi = m_i \psi \text{ , where } i = 1, ..., k.$$

In other words, ψ is the simultaneous eigenvector of the set of eigenvalues

$$m_1, m_2, \ldots, m_k.$$

(This isn't a new requirement, since we already said that the h_i were simultaneously diagonalizable.) It is sometimes convenient to combine these eigenvalues into a vector $\vec{m}$ in k-dimensional space (the dimension of the Cartan subalgebra). This vector is said to be the *weight* of the representation. An N-dimensional representation has a total of N weights, some of which can be identical (i.e., there can be "degeneracy").

In the previous case of $\mathfrak{su}(2)$, the weights were single numbers because the Cartan subalgebra was spanned by a single vector h. In general, *the N weights of an N-dimensional representation are vectors in a k-dimensional vector space*, where k is the dimension of the Cartan subalgebra.

The *roots* of a Lie algebra L are the non-zero weights of the adjoint representation ad_L. Thus we have the following relationship:

$$\{\text{all roots of a Lie algebra}\} \subset \{\text{all weights of a Lie algebra}\}$$

6.4 Tensor Products of Representations

Tensor products are one of the most useful constructions in physics. Using tensor products one can combine representations of groups to produce a wide variety of physically interesting and useful constructions. Let $D^{(1)}(g)_{\alpha\beta}$ and $D^{(2)}(g)_{ab}$ be representations of G acting on vectors $\phi^{(1)}_\beta \in V^{(1)}, \phi^{(2)}_\beta \in V^{(2)}$. We define the tensor product $D^{(1)} \otimes D^{(2)}$ acting on $V^{(1)} \otimes V^{(2)}$ by:

$$(D^{(1)} \otimes D^{(2)})_{\alpha a,\beta b} \equiv D^{(1)}_{\alpha\beta} D^{(2)}_{ab}$$

This acts on $\Phi_{\beta b} \in V^{(1)} \otimes V^{(2)}$ by:

$$\Phi_{\alpha a} \mapsto D^{(1)}_{\alpha\beta} D^{(2)}_{ab} \Phi_{\beta b}$$

A special form of the tensor $\Phi_{\alpha a}$ is the factorized form $\phi^{(1)}_\alpha \phi^{(2)}_a$, but this is not necessary (in other words, not all tensors are the direct product of vectors). For the dimension of the representation, we have:

$$\dim(D^{(1)} \otimes D^{(2)}) = (\dim D^{(1)})(\dim D^{(2)})$$

6.4.1 The Representation of $L(G)$ associated to $D^{(1)} \otimes D^{(2)}$

Let $g \in G$. Set $g = I + tX + \mathcal{O}(t^2)$. Then

$$\begin{aligned} D^{(1)}(g) \otimes D^{(2)}(g) &= (I + td^{(1)}(X)) \otimes (I + td^{(2)}(X)) \\ &= I \otimes I + t(d^{(1)}(X) \otimes I + I \otimes d^{(2)}(X)) + \mathcal{O}(t^2) \end{aligned}$$

Thus the associated representation of $L(G)$ is:

$$d^{(1\otimes 2)} = d^{(1)} \otimes I + I \otimes d^{(2)}.$$

As a result of this form, we see that in the tensor product representation $d^{(1\otimes 2)}$ the eigenvalues of $d^{(1)}$ and $d^{(2)}$ add. Therefore, by definition, the weights add.

6.4.2 Tensor Products of $\mathfrak{su}(2)$ Irreducible Representations

Let j refer to the $d^{(j)}$ spin j irreducible representation of $\mathfrak{su}(2)$. The tensor product $j \otimes j'$ decomposes according to

$$j \otimes j' = (j + j') \oplus (j + j' - 1) \oplus \cdots \oplus |j - j'| \tag{6.1}$$

into a direct sum of irreducible representations. This formula - known as the *Clebsch Gordon series* is used for combining states of particles with spins j and j'. One can verify the formula above by comparing weights on both sides (cf. Example Sheet 3, Problem 1).

Example:
$j = 1$ has weights $\{-1, 0, 1\}$. $j = \frac{1}{2}$ has weights $\{-\frac{1}{2}, +\frac{1}{2}\}$. Thus $1 \otimes \frac{1}{2}$ has weights $\{-\frac{3}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{3}{2}\} = \{-\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}\} \cup \{-\frac{1}{2}, \frac{1}{2}\}$. Thus $1 \otimes \frac{1}{2} = \frac{3}{2} \oplus \frac{1}{2}$.

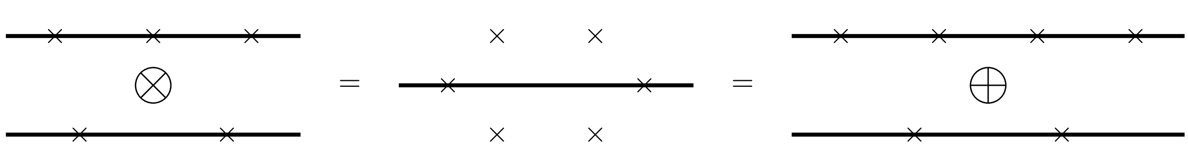


Figure 6.4: A diagrammatic illustration of weights adding.

In Eq 6.1 each state on the right $|J, M\rangle$ with $|j - j'| \leq J \leq j + j'$ can be expressed explicitly as a linear combination of states on the left:

$$|J, M\rangle = \sum_m c_m \, |j, m\rangle \otimes |j', M - m\rangle$$

One writes more formally the coefficients c_m as the Clebsch Gordon coefficients $c^{Jjj'}_{Mmm'}$, which (after a simple rescaling) are also known as the Wigner $3j$ symbols. They are non-vanishing only for $M = m + m'$, which is just that statement that weights add. For more information, see Landau and Lifschitz Quantum Mechanics, §106.

Chapter 7

Gauge Theories

Classical and quantum field theory are constructed from a Lagrangian density $\mathcal{L}$ which is gauge invariant and Lorentz invariant. The action is:

$$S = \int_{\mathbb{R}^4} \mathcal{L}\, d^4x.$$

A gauge theory has a Lie group G acting as a local symmetry, i.e., acting independently at each spacetime point. Fields differing by a gauge transformation are physically the same.

7.1 Scalar Electrodynamics

Here the gauge group G is $U(1)$, which is abelian. This theory describes the electromagnetic field interacting with other electrically charged fields. $U(1)$ is a global symmetry, which lead to charge conservation via Noether's Theorem. Gauge symmetry is local and leads to massless photons (see below).

We begin with the ungauged theory as a complex scalar field $\phi(x)$:

$$\mathcal{L} = \frac{1}{2}\partial_\mu \overline{\phi} \partial^\mu - U(\phi\overline{\phi}),$$

where U us a function (usually a polynomial) of $|\phi|^2 = \overline{\phi}\phi$. We use the convention of $\eta_{\mu\nu} = \mathrm{diag}(+1, -1, -1, -1)$ for the Minkowski metric. So $\partial_\mu \overline{\phi} \partial^\mu \phi = \partial_0 \overline{\phi} \partial_0 \phi - \nabla\phi \cdot \nabla\phi$. This convention gives us a positive kinetic term (time derivatives) in $\mathcal{L}$.

$\mathcal{L}$ is Lorentz invariant and invariant under the global $U(1)$ symmetry $\phi \mapsto \exp(i\alpha)\phi$, $\overline{\phi} \mapsto \exp(-i\alpha)\overline{\phi}$. To obtain a $U(1)$ gauge theory, $\mathcal{L}$ must also be invariant under

$$\phi(x) \mapsto \exp(i\alpha(x))\phi(x), \tag{7.1}$$

where $\alpha(x)$ is an arbitrary (well-behaved) real function. The $U(\overline{\phi}\phi)$ term is already OK as it stands. The derivative terms clearly cause problems with what have so far. The way to solve this is to introduce a new field, the *real gauge potential* $a_\mu(x)$, and *gauge covariant derivative* of ϕ:

$$D_\mu \phi = \partial_\mu \phi - i a_\mu \phi$$

(In this definition of D_μ we have set the coupling $g = 1$) We *postulate* that under Eq 7.1 the gauge potential transforms according to:

$$a_\mu \mapsto a_\mu + \partial_\mu \alpha(x)$$

With these definitions, we see that $D_\mu\phi$ transforms in the same way as ϕ ("covariantly" with ϕ):

$$\begin{aligned} D_\mu\phi \mapsto & \partial_\mu\phi' - ia'_\mu\phi' \\ &= \partial_\mu(e^{i\alpha}\phi) - i(a_\mu + \partial_\mu\alpha)(e^{i\alpha}\phi) \\ &= i(\partial_\mu)e^{i\alpha}\phi + e^{i\alpha}\partial_\mu\phi - ia_\mu e^{i\alpha}\phi - i(\partial_\mu\alpha)e^{i\alpha}\phi \\ &= e^{i\alpha}(\partial_\mu - ia_\mu\phi) \\ &= e^{i\alpha}D_\mu\phi \end{aligned}$$

Similarly, the covariant derivative of $\overline{\phi}$ is $D_\mu\overline{\phi} = \overline{D_\mu\phi} = \partial_\mu\overline{\phi} + ia_\mu\overline{\phi}$, and so

$$\overline{D_\mu\phi} \mapsto \exp(-i\alpha)\overline{D_\mu\phi}$$

under gauge transformation. With these modifications, $D_\mu\phi\overline{D^\mu\phi}$ is gauge invariant and Lorentz invariant. The terms $D_\mu\phi\overline{D^\mu\phi}$ and $U(\overline{\phi}\phi)$ are ingredients in a gauge Lagrangian $\mathcal{L}$.

The field a_μ is also dynamical, and so we need its derivatives in $\mathcal{L}$. The electromagnetic field tensor $f_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu$ is gauge invariant and appears in $\mathcal{L}$. We can easily verify the gauge invariance of $f_{\mu\nu}$:

$$\begin{aligned} f_{\mu\nu} \mapsto & \partial_\mu a'_\nu - \partial_\nu a'_\mu \\ &= \partial_\mu(a_\nu + \partial_\nu\alpha) - \partial_\nu(a_\mu + \partial_\mu\alpha) \\ &= \partial_\mu a_\nu - \partial_\nu a_\mu \text{ , by symmetry of partial derivatives} \\ &= f_{\mu\nu} \end{aligned}$$

We now combine these three ingredients to construct a Lagrangian $\mathcal{L}$ for scalar electrodynamics:

$$\mathcal{L} = -\frac{1}{4}f_{\mu\nu}f^{\mu\nu} + D_\mu\phi\overline{D^\mu\phi} - U(\overline{\phi}\phi).$$

In order to understand the signs, we separate this equation into time and space parts. Recall: $f_{0i} = e_i$ with $\vec{e} = \partial_0\vec{a} - \boldsymbol{\nabla}a_0$ and $b_k = \frac{1}{2}\epsilon_{ijk}f_{ij}$ with $\vec{b} = \boldsymbol{\nabla}\times\vec{a}$, where $\vec{e}$ and $\vec{b}$ are the electric and magnetic fields. Using this notation, we find

$$\begin{aligned} \mathcal{L} &= -\frac{1}{4}(-2\vec{e}\cdot\vec{e} + 2\vec{b}\cdot\vec{b}) + \frac{1}{2}\overline{D_0\phi}D^0\phi - \frac{1}{2}\overline{\vec{D}\phi}\cdot\vec{D}\phi - U(\overline{\phi}\phi) \\ &= \underbrace{\frac{1}{2}\vec{e}\cdot\vec{e} + \frac{1}{2}\overline{D_0\phi}D^0\phi}_{\text{“Kinetic Terms”}} \underbrace{-\frac{1}{2}\vec{b}\cdot\vec{b} - \frac{1}{2}\overline{\vec{D}\phi}\cdot\vec{D}\phi - U(\overline{\phi}\phi)}_{\text{“Potential Terms”}} \end{aligned}$$

This is a "natural Lagrangian" in the sense that it takes the form $L = T - V$, where T is quadratic in time derivatives. However, a_0 is *not* dynamical. For completeness, we give the Euler-Lagrangian equations, though their derivation is not a part of this course:

$$\begin{aligned} D_\mu D^\mu\phi &= -2U'(\overline{\phi}\phi)\phi \\ \partial_\mu f^{\mu\nu} &= -\frac{i}{2}(\overline{\phi}D^\nu\phi - \phi\overline{D^\nu\phi}), \end{aligned}$$

where U' is the derivative of U with respect to its argument $\overline{\phi}\phi$.

Note: These are second-order evolution equations for $\phi, \vec{a}$. The $\nu = 0$ component is rather different (think: Gauss' law...).

7.1.1 Field Tensor from Covariant Derivatives

$$\begin{aligned}
[D_\mu, D_\nu]\phi &= (\partial_\mu - ia_\mu)(\partial_\nu - ia_\nu)\phi - (\mu \leftrightarrow \nu) \\
&= (\partial_\mu - ia_\mu)(\partial_\nu\phi - ia_\nu\phi) - (\mu \leftrightarrow \nu) \\
&= \cancel{\partial_\mu\partial_\nu\phi} - i(\partial_\mu a_\nu)\phi - \cancel{a_\nu\partial_\mu\phi} - \cancel{a_\mu\partial_\nu\phi} - \cancel{a_\mu a_\nu\phi} - (\mu \leftrightarrow \nu) \\
&= -i(\partial_\mu a_\nu - \partial_\nu a_\mu)\phi \\
&= -if_{\mu\nu}\phi
\end{aligned}$$

So we see that the field tensor arises as a commutator of covariant derivatives.

7.2 A Non-Abelian Gauge Theory: Scalar Yang-Mills Theory

We now extend our formalism to a general gauge group G. We fix $G = U(n)$ or some Lie subgroup thereof (e.g., $SU(n), SO(n), \dots$). We introduce the field

$$\Phi(x) = \begin{pmatrix} \Phi_1(x) \\ \vdots \\ \Phi_n(x) \end{pmatrix},$$

which has n components. As before, we can consider the global action of G

$$\Phi(x) \mapsto g\Phi(x),$$

where $g \in G$ is independent of x. However, we will require our theory to be invariant under gauge transformations

$$\Phi(x) \to g(x)\Phi(x),$$

where $g \in G$ *also depends on location.* As previously, this inspires us to introduce the covariant derivative

$$D_\mu\Phi = (\partial_\mu + A_\mu)\Phi,$$

with $A_\mu(x) \in L(G)$. In particular, A_μ (as a member of the Lie algebra of $U(N)$) is an $n \times n$ antihermitian matrix. Although there are certain *geometric motivations*, we *postulate* that A_μ gauge transforms as:

$$A_\mu \mapsto A'_\mu = gA_\mu g^{-1} - (\partial_\mu g)g^{-1}$$

Remark. Earlier in the notes (see the discussion leading to (3.2)) we discuss why terms of the form $g^{-1}(\partial_\mu g)$ and $(\partial_\mu g)g^{-1}$ are also elements of the Lie algebra $L(G)$. In our discussion of the adjoint representation of G, we found that $\mathrm{Ad}_g X = gXg^{-1} \in L(G)$ where $g \in G$ and $X \in L(G)$.

Remark. This new definition for the transformation of the gauge field A_μ is consistent with our previous postulate for the transformation law in the abelian case of scalar electrodynamics. In other words, $ga_\mu g^{-1} - (\partial_\mu g)g^{-1} = a_\mu - i\partial_\mu\alpha(x)$ when $g = e^{i\alpha(x)}$. Since the covariant derivative is defined slightly differently in the non-abelian case, the appearance of i is not a particular source of worry.

We see that $D_\mu\Phi$ transforms covariantly with Φ:

$$\begin{aligned}
D_\mu\Phi = (\partial_\mu + A_\mu)\Phi &\mapsto (\partial_\mu + gA_\mu g^{-1} - (\partial_\mu g)g^{-1})g\Phi \\
&= \cancel{(\partial_\mu g)\Phi} + g(\partial_\mu\Phi) + gA_\mu\Phi - \cancel{(\partial_\mu g)\Phi} \\
&= g(\partial_\mu\Phi + A_\mu\Phi) \\
&= gD_\mu\Phi
\end{aligned}$$

We now look for the *Yang-Mills field tensor* by examining the commutator of covariant derivatives:

$$\begin{aligned}
[D_\mu, D_\nu]\Phi &= (\partial_\mu + A_\mu)(\partial_\nu + A_\nu)\phi - (\mu \leftrightarrow \nu) \\
&= \partial_\mu\partial_\nu\phi + A_\mu\partial_\nu\Phi + \partial_\mu(A_\nu\Phi) + A_\mu A_\nu\Phi - (\mu \leftrightarrow \nu) \\
&= \cancel{\partial_\mu\partial_\nu\phi} + \cancel{A_\mu\partial_\nu\Phi} + \partial_\mu A_\nu\Phi + \cancel{A_\nu\partial_\mu\Phi} + A_\mu A_\nu\Phi - (\mu \leftrightarrow \nu) \\
&= (\partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu])\Phi \\
&= F_{\mu\nu}\Phi,
\end{aligned}$$

where $F_{\mu\nu} \equiv (\partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu])$ is the field tensor.

Note: The final term can be interpreted abstractly as the Lie bracket and not simply as the matrix commutator.
Note: $F_{\mu\nu} \in L(G)$, $F_{\mu\nu} = -F_{\nu\mu}$.

Proposition 3. *Under gauge transformations, $F_{\mu\nu} \mapsto gF_{\mu\nu}g^{-1}$*

Proof. $[D_\mu, D_\nu]\Phi \mapsto g[D_\mu, D_\nu]\Phi = gF_{\mu\nu}\Phi = (gF_{\mu\nu}g^{-1})(g\Phi)$. Thus we see that $F_{\mu\nu} \mapsto gF_{\mu\nu}g^{-1}$. □

Remark. One can also compute this directly in terms of the behavior of A_μ, but the calculation is roughly half a page long instead of simply a line.

7.2.1 Lagrangian Density

For the Yang-Mills gauge potential A_μ coupled to the scalar Φ, the Lagrangian density is given by:

$$\mathcal{L} = \frac{1}{4}\operatorname{Tr}(F_{\mu\nu}F^{\mu\nu}) + \frac{1}{2}(D_\mu\Phi)^\dagger D^\mu\Phi - U(\Phi^\dagger\Phi).$$

Note that there's nothing strange happening with the trace operator, since $F_{\mu\nu}F^{\mu\nu} \in L(G)$. Lorentz invariance of $\mathcal{L}$ is clear, since all the Lorentz indices are contracted. However we still need to check gauge invariance:

$$\begin{aligned}
&\Phi \mapsto g\Phi \\
&\Phi^\dagger \mapsto \Phi^\dagger g^\dagger \text{since we're in U(n)} \\
\therefore\ &\Phi^\dagger\Phi \mapsto \Phi^\dagger g^{-1}g\Phi = \Phi^\dagger\Phi,
\end{aligned}$$

so $\Phi^\dagger\Phi$ is invariant. Since $D_\mu\Phi$ transforms as Φ, the same argument applies to the second term in the Lagrangian. For the final terms we see:

$$\begin{aligned}
\operatorname{Tr}(F_{\mu\nu}F^{\mu\nu}) &\mapsto \operatorname{Tr}(gF_{\mu\nu}g^{-1}gF^{\mu\nu}g^{-1}) \\
&= \operatorname{Tr}(F_{\mu\nu}IF^{\mu\nu}gg^{-1}) \text{ , since Tr is invariant under cyclic permutations} \\
&= \operatorname{Tr}(F_{\mu\nu}F^{\mu\nu})
\end{aligned}$$

Thus, the Lagrangian is gauge invariant.

As before, we have the covariant derivative $D_\mu\Phi = (\partial_\mu + A_\mu)\Phi$ and the field tensor $F_{\mu\nu} = (\partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu])$. We can look at the kinetic terms in the Lagrangian:

$$-\frac{1}{4}\operatorname{Tr}(F_{0i}F_{0i}) + \frac{1}{2}(D_\mu\Phi)^\dagger D^\mu\Phi$$

Note: F_{0i} is an anti-hermitian matrix, so $-\frac{1}{4}\operatorname{Tr}(F_{0i}F_{0i})$ is positive.
Why is this true? If A is anti-hermitian,

$$\operatorname{Tr} A^2 = A_{\alpha\beta}A_{\beta\alpha} = -A_{\alpha\beta}A^*_{\alpha\beta} = -\sum_{\alpha,\beta}|A_{\alpha\beta}|^2.$$

Hence all kinetic terms are positive. This is a consequence of the gauge group G being unitary (so elements of its Lie algebra are anti-hermitian).

7.2.2 Adjoint Covariant Derivative

We remark that not all fields transform as $\Phi \mapsto g\Phi$ under gauge transformations. We could also have a scalar field $\Psi \in L(G)$ transforming as $\Psi \mapsto g\Psi g^{-1}$ (This looks like the transformation of $F_{\mu\nu}$). In this case, the covariant derivative is:

$$D_\mu\Psi = \partial_\mu\Psi + [A_\mu, \Psi].$$

To check this result, one substitutes the transformation laws $\Psi \mapsto g\Psi g^{-1}$ and $A_\mu \mapsto A'_\mu = gA_\mu g^{-1} - (\partial_\mu g)g^{-1}$ for the scalar and gauge fields, respectively, into the proposed formula. One then sees after several lines of computation that the covariant derivative transforms as Ψ, i.e., $D_\mu\Psi \mapsto g(D_\mu\Psi)g^{-1}$.

7.2.3 General Covariant Derivative

Abstractly, we write $D_\mu = \partial_\mu + A_\mu$, where $A_\mu \in L(G)$ and $[D_\mu, D_\nu] = F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$. In order to act concretely, we need a field $\Phi \in V$ (V is a vector space) and a representation $\mathcal{D}$ of G. (Note that since D in this chapter refers to the gauge covariant derivative we have adopted the script $\mathcal{D}$ to stand for a representation.) Then under gauge transformations the field transforms as $\Phi \mapsto \mathcal{D}(g)\Phi$ and $D_\mu\Phi = \partial_\mu\Phi + d(A_\mu)\Phi$, where d is the representation of $L(G)$ associated to $\mathcal{D}$.

Claim: $D_\mu\Phi$ transforms like Φ under gauge transformations.

Proof. We know $\Phi \mapsto \mathcal{D}(g)\Phi$ and $d(A_\mu) \mapsto \mathcal{D}(g)d(A_\mu)\mathcal{D}(g^{-1}) - (\partial_\mu\mathcal{D}(g))\mathcal{D}(g^{-1})$ Thus:

$$\begin{aligned} D_\mu\Phi = \partial_\mu\Phi + d(A_\mu)\Phi &\mapsto \partial_\mu(\mathcal{D}(g)\Phi) + (\mathcal{D}(g)d(A_\mu)\mathcal{D}(g^{-1}) - (\partial_\mu\mathcal{D}(g))\mathcal{D}(g^{-1}))\mathcal{D}(g)\Phi \\ &= \cancel{(\partial_\mu\mathcal{D}(g))\Phi} + \mathcal{D}(g)(\partial_\mu\Phi) + \mathcal{D}(g)d(A_\mu)\Phi - \cancel{(\partial_\mu\mathcal{D}(g))\Phi} \\ &= \mathcal{D}(g)(\partial_\mu\Phi + d(A_\mu)\Phi) \\ &= \mathcal{D}(g)D_\mu\Phi \end{aligned}$$

Thus D_μ transforms "covariantly" with Φ under gauge transformations. □

7.2.4 The Field Equation of Pure Yang-Mills Theory

The field equation fo pure Yang-Mills theory is:

$$\partial_\mu F^{\mu\nu} + [A_\mu, F^{\mu\nu}] = 0$$

Remark: This is an equation for A_μ involving second-order derivatives. In Yang-Mills theory, $F^{\mu\nu}$ is a derived quantity, while A_μ is considered fundamental (contrast with electromagnetism). We may rewrite the equation of motion more compactly using the adjoint covariant derivative:

$$D_\mu F^{\mu\nu} = 0.$$

7.2.5 Classical Vacuum

Naively, the classical vacuum is $A_\mu = 0$. We can gauge transform this to $A_\mu = -(\partial_\mu g)g^{-1}$. This is a "pure gauge," which is physically the same as the vacuum.
Claim: The Yang-Mills field tensor vanishes in both the classical vacuum and the pure gauge transformed vacuum: $F^{\mu\nu} = 0$ in both A_μ and $A_\mu = -(\partial_\mu g)g^{-1}$

Proof.

$$\begin{aligned}
F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \\
&= \partial_\mu(-(\partial_\nu g)g^{-1}) - \partial_\nu(-(\partial_\mu g)g^{-1}) + [-(\partial_\mu g)g^{-1}, -(\partial_\nu g)g^{-1}] \\
&= \cancel{-(\partial_\mu\partial_\nu g)g^{-1}} - (\partial_\nu g)(\partial_\mu g^{-1}) + \cancel{(\partial_\nu\partial_\mu g)g^{-1}} + (\partial_\mu g)(\partial_\nu g^{-1}) + (\partial_\mu g)g^{-1}(\partial_\nu g)g^{-1} - (\partial_\nu g)g^{-1}(\partial_\mu g)g^{-1} \\
&= (\partial_\mu g)[(\partial_\nu g^{-1}) + g^{-1}(\partial_\nu)g^{-1}] - (\partial_\nu g)[(\partial_\mu g^{-1}) + g^{-1}(\partial_\mu)g^{-1}] \\
&= (\partial_\mu g)[\partial_\nu(g^{-1}g)] - (\partial_\nu g)[\partial_\mu(g^{-1}g)]
\end{aligned}$$

But $g^{-1}g = 1 = \text{const}$, so $\partial_\nu(g^{-1}g) = 0$. Thus $F^{\mu\nu} = 0$ in the pure gauge $A_\mu = -(\partial_\mu g)g^{-1}$. □

If there is a standard scalar field Φ as well, the $\Phi = 0$ and $D_\mu\Phi = 0$ (provided $\Phi = 0$ minimizes the potential $U(\Phi^\dagger\Phi)$). The quantized particles of Yang-Mills theory are massless but confined. They are called gluons in the $SU(3)$ theory of quantum chromodynamics. This phenomenon is only partially understood.

7.3 A Very Brief Introduction to Mass and the Higgs Mechanism

Recall the Klein-Gordon theory for a scalar field ϕ is governed by the Lagrangian:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2.$$

As we derived several times in quantum field theory, the equation of motion is of course the Klein-Gordon equation:

$$\partial_\mu\partial^\mu\phi + m^2\phi = 0.$$

This equation has wavelike solutions of the form $\phi = \exp[i(\omega t - \vec{k}\cdot\vec{x})]$ provided $\omega^2 = |\vec{k}|^2 + m^2$. Here m is the "mass parameter" of the field. Particles arise when the field is quantized. Particle energy is given by $E = \hbar\omega$; particle momentum is given by $\vec{p} = \hbar\vec{k}$. Thus $E^2 = |\vec{p}|^2 + (\hbar m)^2$. So in units were $c = 1$, a particle has mass $\hbar m$.

7.3.1 Electrodynamics

We have the Lagrangian $\mathcal{L} = -\frac{1}{4}f_{\mu\nu}f^{\mu\nu}$ with $f_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu$. The Maxwell equation says:

$$\partial_\mu f^{\mu\nu} = 0 \implies \partial_\mu(\partial^\mu a^\nu) - \partial^\nu(\partial_\mu a^\mu) = 0.$$

If we work in the Coulomb gauge, $\partial_i a_i = 0, i = 1, 2, 3$,i.e., $\boldsymbol{\nabla}\vec{a} = 0$. Let us take a closer look at the $\nu = 0$ equation: $\partial_\mu a^\mu = \partial_0 a_0$. Thus $\partial_\mu\partial^\mu a^0 - \partial^0(\partial_0 a_0) = (\partial^0\partial_0 - \partial_i\partial^i)a^0 - \partial^0\partial_0 a^0 = 0 \implies \boldsymbol{\nabla}^2 a_0 = 0$. Without any special boundary conditions or sources, the solution is evidently $a_0 = 0$. The remaining equations are $(\partial_\mu\partial^\mu)a_i = 0$. These are *three massless wave equations.*

We see that polarization is transverse to momentum: If $a_i = \epsilon_i \exp[i(\omega t - \vec{k}\cdot\vec{x})]$, then $\omega^2 = |\vec{k}|^2$ (massless!). The gauge condition $\partial_i a_i = 0$ tells us $\vec{k}\cdot\vec{\epsilon} = 0$. After quantization, the *photon* has the following properties:

- zero mass, so energy satisfies $E = |\vec{p}|$
- 2 polarizations states: $\vec{\epsilon} \perp \vec{p}$
- non-vanishing momentum: $\vec{p} = 0$

7.3.2 Perturbative Effect of Interaction of EM Field with a Charged Scalar Field

The scalar particle mass m is renormalized, thereby changing its value. The photon, however, remains massless. This is a true result, but it is complicated to prove. However, it may be understood as arising from two facts:

(a) No term of the form $\frac{1}{2}M^2 a_\mu a^\mu$ is allowed, because it wouldn't be gauge invariant

(b) A massive vector particle can have $\vec{p} = 0$, and it then has polarization states (i.e., $\vec{\epsilon} \cdot \vec{p} = 0$ doesn't constrain $\vec{\epsilon}$ if $\vec{p} = 0$, and a continuous perturbation cannot change the number of polarization states.

The *Higgs mechanism* evades these two previous objections. In scalar electrodynamics, the "photon" can become massive. The abelian Higgs model has the Langrangian:

$$\mathcal{L} = -\frac{1}{4} f_{\mu\nu} f^{\mu\nu} + \frac{1}{2}\overline{D_\mu \phi} D^\mu \phi + \frac{1}{2}\mu^2 \overline{\phi}\phi - \frac{1}{4}\lambda(\overline{\phi}\phi)^2. \tag{7.2}$$

Evidently, the potential is $U = -\frac{1}{2}\mu^2\overline{\phi}\phi + \frac{1}{4}\lambda(\overline{\phi}\phi)^2$.

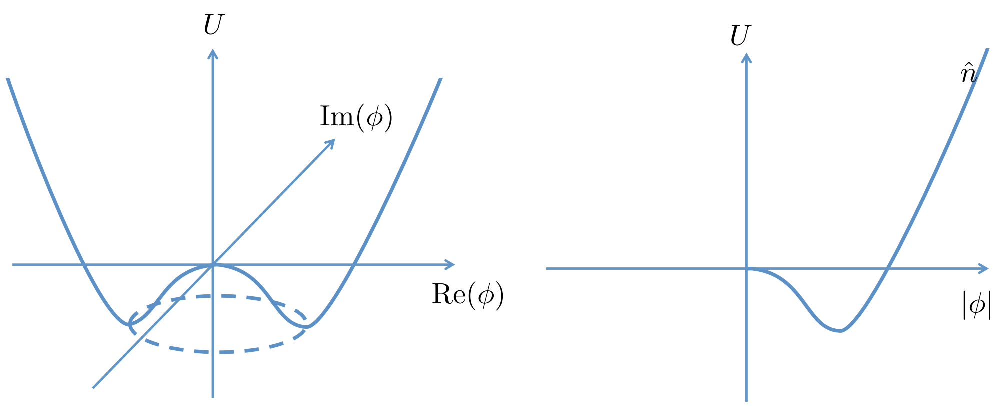


Figure 7.1: The so-called "Mexican hat potential"

Often it is convenient to shift the potential up and write $U = \frac{1}{4}\lambda(|\phi|^2 - v^2)^2$, where $\mu^2 = \lambda v^2$. This shift is by a constant value of $\frac{1}{4}\lambda v^4$, which has no effect on the field equations. After this shift, the vacuum is *not* at $\phi = 0$, but where $|\phi| = v$. Further, the vacuum is *degenerate* (related by gauge transformations). Moreover, the vacuum is a *non-trivial orbit* of the gauge group $U(1)$. We call v the *vacuum expectation value* of ϕ. (Note that this "quantum language is a bit of leap in the dark, since we're still working completely classically). The simplest vacuum is $\phi = v, a_\mu = 0$. We can gauge transform this vacuum to $\phi = v \exp i\alpha(x), a_\mu = \partial_\mu$. A natural question is then how this effects the two arguments above that kept the photon massless.

(a') There is now a mass term for a_μ:

Part of $\frac{1}{2}\overline{D_\mu\phi}D^\mu\phi$ is $\frac{1}{2}(\overline{-ia_\mu\phi})(-ia^\mu\phi) = \frac{1}{2}a_\mu a^\mu|\phi|^2$. Close to the vacuum, this is $\frac{1}{2}v^2 a_\mu a^\mu +$ higher order terms and dominates in the equation for a_μ. Thus the "photon" acquires mass v and gets the new name *massive vector boson.*

Remark: The term we've singled out would, by itself, violate gauge invariance. However, the other terms present in the Lagrangian preserve the gauge invariance.

(b') The massive vector boson has *three* polarization states. We can impose the gauge condition

$$\text{Im } \phi = 0, \phi(x) = v + \eta(x),$$

where $\eta(x)$ real. We cannot simultaneously impose the Coulomb gauge. Thus the physical particles are:

- massive vector bosons with three states, or
- a real scalar particle H from the field η with 1 state.

Alternatively, in the coulomb gauge, we need to allow $\phi(x) = \exp[i\beta(x)](v + \eta(x))$ with $\beta(x), \eta(x)$ real and the third polarization state of the vector boson coming from β. For the *mass of the Higgs particle H*, look at quadratic terms in η (with $a_\mu = 0$).

$$\mathcal{L}_\eta = \frac{1}{2}\partial_\mu\eta\partial^\mu\eta - \frac{1}{4}\lambda((v+\eta)^2 - v^2)^2 = \frac{1}{2}\partial_\mu\eta\partial^\mu\eta - \lambda v^2\eta^2 + \dots,$$

which we find from substituting $\phi(x) = v + \eta(x)$ into the Lagrangian of the abelian Higgs model, (7.2). Thus we see that $m_H = \sqrt{2\lambda}v = \sqrt{2}\mu$. (Here we've omitted the coupling e). The main result is: *The masses of the vector bosons and Higgs particle are unrelated, so we had to measure the Higgs mass experimentally.*

7.3.3 Higgs Mechanism in Nonabelian Case

The standard model as a $U(2)$ electroweak gauge group and scalar field Φ, which is a doublet of complex scalar fields. ($U(2)$ acts by the fundamental representation). The Higgs mechanism occurs if $\Phi \neq 0$ in the classical vacuum, i.e., if the minimum of the potential $U(\Phi)$ is on a *non-trivial orbit* of the gauge group G. We expect the orbit to have the form G/K for some subgroup $K \leq G$. Then the symmetry of G is broken, while for K it remains unbroken. Massless gauge bosons remain for K. In the standard model, $G = U(2), K = U(1)$, and we get:

- 3 massive vector bosons W^+, W^-, Z
- 1 massless boson, the photon

Chapter 8

Quadratic Forms on Lie Algebras and the Geometry of Lie Groups

8.1 Invariant Quadratic Forms

For a matrix Lie algebra $L(G)$, we have the quadratic form $L(G) \times L(G) \to L(G)$ (symmetric innter product) given by:

$$(X, Y) = \mathrm{Tr}(XY),$$

where $X, Y \in L(G)$. This quadratic form is invariant in the sense following sense:

$$(X, Y) \mapsto (gXg^{-1}, gYg^{-1}) = \mathrm{Tr}(gXg^{-1}gYg^{-1}) = \mathrm{Tr}(XY) = (X, Y),$$

where $g \in G$. One can also use a different representation d of $L(G)$ and define $(X, Y)_d = \mathrm{Tr}(d(X), d(Y))$, which is similarly invariant: $(gXg^{-1}, gYg^{-1})_d = (X, Y)_d$. This is a result of the following lemma:

Lemma 8.1.1. *If d is a representation of $L(G)$ associated to the representation D of G, then $d(gXg^{-1}) = D(g)d(X)D(g^{-1})$.*

Proof. Let D and d be as described above. Let g be a curve in G of the form $g(t) = I + Yt + \mathcal{O}(t^2)$, $Y \in L(G)$. Clearly we also have $g^{-1}(t) = I - Yt + \mathcal{O}(t^2)$.

$$\begin{aligned}
g(t)Xg^{-1}(t) &= (I + Yt)X(I - Yt) + \mathcal{O}(t^2) \\
&= (X + YXt)(1 - Yt) + \mathcal{O}(t^2) \\
&= X - YXt + YXt + \mathcal{O}(t^2) \\
&= X + [Y, X]t \\
\implies d(g(t)Xg^{-1}(t)) &= d(X + [Y, X]t) \\
&= d(X) + d([Y, X])t + \mathcal{O}(t^2)
\end{aligned}$$

By definition, $D(g) = I + d(Y)t + \mathcal{O}(t^2)$ and $D(g^{-1}) = I - d(Y)t + \mathcal{O}(t^2)$. Thus:

$$\begin{aligned}
D(g)d(X)D(g^{-1}) &= (I + d(Y)t + \mathcal{O}(t^2))d(X)(I - d(Y)t + \mathcal{O}(t^2)) \\
&= (d(X) + d(Y)d(X)t)(1 - d(Y)t) + \mathcal{O}(t^2) \\
&= d(X) - (d(X)d(Y) - d(Y)d(X))t + \mathcal{O}(t^2) \\
&= d(X) + [d(Y), d(X)] + \mathcal{O}(t^2)
\end{aligned}$$

But d commutes with the bracket $[\ ,\ ]$, so $d([Y,X]) = [d(Y), d(X)]$. Thus $d(gXg^{-1}) = D(g)d(X)D(g^{-1})$. □

As a special case, we consider the adjoint representation of $L(G)$, which gives rise to the so-called *Killing form*:

$$K(X,Y) = \mathrm{Tr}(\mathrm{ad}_X, \mathrm{ad}_Y)$$

The Killing form is the most basic inner product, because every Lie algebra has an adjoint representation. In practices, these inner products are usually related. They differ by a constant multiple if $L(G)$ is "simple" (to be made precise soon).
The infinitesimal version of the invariance above is the *associative property*:
Claim: $(X,[Y,Z])_d = ([X,Y],Z)_d$

Proof. (For the intial inner product in this section)
$\mathrm{Tr}(X,[Y,Z]) = \mathrm{Tr}(XYZ) - \mathrm{Tr}(XZY) = \mathrm{Tr}(XYZ) - \mathrm{Tr}(YXZ) = \mathrm{Tr}([X,Y]Z)$ □

What does this have to do with an infinitesimal version of the invariance under the conjugation $X \mapsto gXg^{-1}$? If we let $g = I + tZ \in G$ with t infinitesimal, then we have:

$$\begin{aligned}
(X,Y) &= ((1+tZ)X(1-tZ), (1+tZ)Y(1-tZ)) \\
&= \mathrm{Tr}((X+tZX)(1-tZ)(Y+tZY)(1-tZ)) \\
&\ \vdots \\
&= \mathrm{Tr}(XY + tZXY - tZYX) \\
&= \mathrm{Tr}(XY) + \mathrm{Tr}(Y,[tZ,X]) - Tr([Y,tZ],X) \\
\implies & \mathrm{Tr}(Y,[tZ,X]) = \mathrm{Tr}([Y,tZ],X),
\end{aligned}$$

which is just the associative property.

8.2 Non-Degeneracy of the Killing Form

Let $\{T_i\}$ be a basis for $L(G)$. The Killing form becomes a symmetric matrix:

$$\kappa_{ij} = K(T_i, T_j) = \mathrm{Tr}(\mathrm{ad}_{T_i}, \mathrm{ad}_{T_j}).$$

Now ad_{T_i} is defined by $\mathrm{ad}_{T_i} X = [T_i, X]$, where $X \in L(G)$. Let $X = X_l T_l$ (basis expansion). Then

$$\mathrm{ad}_{T_i} X = [T_i, X_l T_l] = X_l [T_i, T_l] = c_{ilk} X_l T_k \implies (\mathrm{ad}_{T_i})_{kl} = c_{ilk}$$

Although the indices looks perhaps a bit strange in the final equality of the last line, they are in fact correct. Therefore we have $\mathrm{Tr}(\mathrm{ad}_{T_i} \mathrm{ad}_{T_j}) = (\mathrm{ad}_{T_i})_{kl} (\mathrm{ad}_{T_j})_{lk} = c_{ilk} c_{jkl}$. So the $(ij)^{\text{th}}$ component of the Killing form is given by:

$$\kappa_{ij} = c_{ilk} c_{jkl}.$$

Example. For the standard basis of $\mathfrak{su}(2)$, $\kappa_{ab} = \epsilon_{adc}\epsilon_{bcd} = -2\delta_{ab}$. This is a non-degenerate Killing form.

Definition 27. The Killing form κ_{ij} is *non-degenerate* if all of its eigenvalues are non-zero. Equivalently, $\det \kappa_{ij} \neq 0$, since $\det = \Pi\lambda_i$, where the λ_i are the eigenvalues.

Definition 28. A Lie algebra L(G) is said to be *semi-simple* if its Killing form is non-degenerate.

Theorem 8.2.1. *A semi-simple Lie algebra $L(G)$ has a decomposition into mutually commuting simple factors $L(G_i)$ such that*

$$L(G) = L(G_1) \otimes L(G_2) \otimes \cdots \otimes L(G_k),$$

where the $L(G_i)$ are semi-simple and cannot be reduced further.

Proof. Omitted (but see Example Sheet 4, Problem 6 for some related details). The general idea is that a non-degenerate Killing form gives a way to build "orthogonal complements," which leads to a direct sum decomposition of the vector space. □

Fact: In the case of $L(G)$ semi-simple, the associated group G has the structure

$$G = G_1 \times G_2 \times \cdots \times G_k / \{\text{discrete group}\}.$$

8.3 Compactness

If κ_{ij} is *negative definite*, $L(G)$ is said to be of *compact type*. G is then a *compact group* (as a topological manifold). For example, $SU(n)$ is simple and compact. (Note: A simple Lie group is a connected, non-abelian Lie group G without non-trivial connected normal subgroups.) In the case of a compact type, we can find an *adapted basis* $\{T_i\}$ such that $\kappa_{ij} = -\mu\delta_{ij}$, for some constant $\mu > 0$.

Then $K(T_i, [T_j, T_k]) = K(T_i, c_{jkl}T_l) = c_{jkl}K_{il} = -\mu c_{jki}$, and $K([T_i, T_j], T_k) = K(c_{ijl}T_l, T_k) = c_{ijl}K_{lk} = -\mu c_{ijk}$. Thus by the associative property we have:

$$c_{ijk} = c_{jki} = c_{kij} = -c_{jik} = -c_{kji} = -c_{ikj}.$$

In other words, *the structure constants are totally antisymmetric in the adapted basis.*

8.4 Universal Enveloping Algebra

In a matrix Lie algebra $L(G)$, we can deal with products XY (which isn't necessary closed in $L(G)$) as well as products $[X, Y]$ (which *is* closed in $L(G)$). In an abstract Lie algebra L, we can define the *universal enveloping algebra* (UEA) to be the formal span of $\{I, L, L \otimes L, L \otimes L \otimes L, \dots\}$, allowing sums and products of elements of L. The UEA is subject to the one rule: $XY - YX = [X, Y]$, as defined in L. In the UEA we see that there are new identities:

(a)

$$\begin{aligned}[X, YZ] &= XYZ - YZX = XYZ - YXZ + YXZ - YZX \\ &= [X, Y]Z + Y[X, Z]\end{aligned}$$

(b)

$$\exp X = I + X + \frac{1}{2}X^2 + \cdots \in \text{UEA}$$

Thus The connected component of G is in the UEA of $L(G)$. (Note: This is a bit of a white lie, as the UEA technically only contains *finite* sums, while the definition of exp involves infinite sums. However, we assume that the idea may be extended in a more careful treatment to include some sort of "formal completion of the UEA" without changing how we view the objects in practice.)

There is also an analogy with the Dirac matrix algebra, which is the span of $\{I, \gamma^\mu, \gamma^\mu\gamma^\nu, \gamma^\mu\gamma^\nu\gamma^\sigma, ...\}$ subject to the constrain $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}I$.

8.5 Casimir Operators

Let $L(G)$ be a Lie algebra of simple (irreducible in the sense that it isn't a direct sum of subspaces) compact (κ_{ij} negative definite) type with adapted basis $\{T_i\}$. Then $\kappa_{ij} = -\delta_{ij}$ (here the basis vectors are normalized so that $\mu = 1$.

Definition 29. The *Casimir operator* in the UEA is $C = \sum_i T_i T_i$.

Lemma 8.5.1. $[X, C] = 0$ *for all* $X \in L(G)$*,i.e., the Casimir commutes with every element in the Lie algebra.*

Proof. It is sufficient to set X equal to the basis vector T_j.

$$\begin{aligned}
[T_j, C] &= \sum_i [T_j, T_i T_i] \\
&= \sum_i (T_i [T_j, T_i] + [T_j, T_i] T_i) \\
&= \sum_i (T_i c_{jik} T_k + c_{jik} T_k T_i) \\
&= \sum_i c_{jik} (T_i T_k + T_k T_i) = 0,
\end{aligned}$$

where the last line follows because the antisymmetric structure constant (adapted basis) is contracted against a symmetric quantity. □

Thus the Casimir C is a central element (commutes with everything) in the UEA. However, this does not mean that it is proportional to 1. Consider now C in a representation d of $L(G)$ of dimension N:

$$C_d = \sum_i d(T_i) d(T_i),$$

where C_d is now simply an $n \times n$ matrix. By the same proof as above, $[d(X), C_d] = 0, \forall X \in L(G)$.

Claim: If d is irreducible, then by Schur's lemma $C_d = c_d I_N$.

Proof. Example Sheet 3, Exercise 7 □

Here the Casimir eigenvalue c_d is a useful way to characterize the irreducible representation d, since c_d depends on the irreducible representation d.

Example. Consider the fundamental ($j = \frac{1}{2}$) representation of $\mathfrak{su}(2)$. The adapted basis is $T_a = -\frac{i}{2\sqrt{2}}\sigma_a$ (notice the factor of $\sqrt{2}$). Then:

$$C_{j=\frac{1}{2}} = \sum_a (T_a)^2 = -\frac{1}{8} \sum_a \sigma_a^2 = -\frac{3}{8}.$$

Thus $c_{j=\frac{1}{2}} = -\frac{3}{8}$.

Example. Consider the well-known angular momentum operators J_1, J_2, J_3 from physics. They have the Casimir operator:

$$C = (-\frac{1}{2})\vec{J}^2 = (-\frac{1}{2})(J_1^2 + J_2^2 + J_3^2)$$

Quoting the familiar result, the eigenvalue in the spin j irreducible representation is $c_{(j)} = -\frac{1}{2}j(j+1)$. Muenter's *Quantentheorie*, for example, has for more information in the context of quantum mechanics.

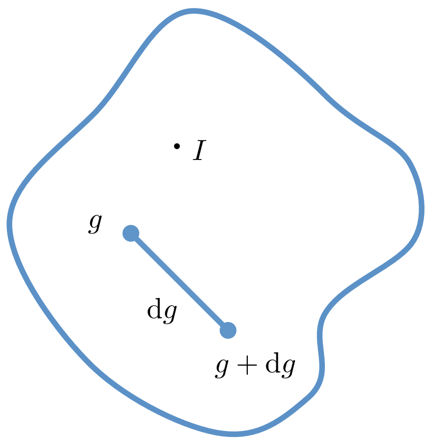
8.6 Metric on G

(Note: In what follows the assumption of a simple Lie group is non essential, but with this assumption the Riemannian metric we find is unique up to a constant multiple. The assumption of compactness gives a Riemannian metric. Otherwise we can wind up with a Lorentzian metric, for example.)

Suppose G is a simple matrix Lie group of compact type. Let $X, Y \in L(G)$ be tangent vectors at I. Define $(X, Y) = -\operatorname{Tr}(XY)$. Up to a constant multiple, this is the unique positive definite quadratic form on $L(G)$ which is also invariant under the action of conjugation on X, Y by G. We can use this quadratic form to define a Riemannian metric on G:

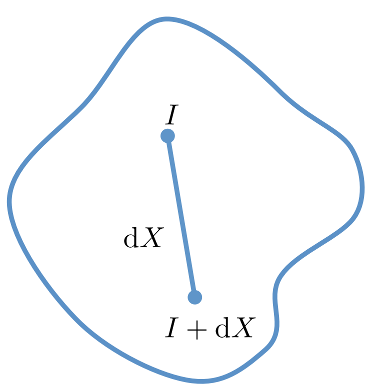
$$ds^2 = -\operatorname{Tr}\left((dg)g^{-1}(dg)g^{-1}\right), \tag{8.1}$$

where dg is an infinitesimal tangent element to G at g. By the argument used previously leading to (3.2), we see that $(dg)g^{-1} \in L(G)$. Interpretation: ds^2 is the squared length of dg.



Note: We must use g^{-1} to map dg back to the identity, where we can use the Lie algebra inner product that we already have.

If we're near the identity, $g = I$ and $g + dg = I + dX$, where dX is an infinitesimal element of $L(G)$. Then $(dg)g^{-1} = dX$ and $ds^2 = -\operatorname{Tr}(dXdX)$.



A natural question to ask is how our inner product behaves under translations by group elements.

(a) Consider left translations by g_0 on dX. We send $I \mapsto g_0$ and $I + dX \mapsto g_0 + g_0 dX$. Then:

$$(dg)g^{-1} \mapsto (g_0 dX)g_0^{-1},$$

which implies

$$ds^2 \mapsto -\operatorname{Tr}(g_0 dX g_0^{-1} g_0 dX g_0^{-1}) = -\operatorname{Tr}(dX dX).$$

So the metric is invariant under right and left translations.

(b) Consider the right translation: $I \mapsto g_0$ and $I + dX \mapsto g_0 + dX g_0$. Then:

$$(dg)g^{-1} \mapsto (dX g_0)g_0^{-1} = dX \implies ds^2 \mapsto -\operatorname{Tr}(dX dX)$$

Thus we've found that our metric is invariant under the action of $G \times G$ on G; such metric is sometimes called a *bi-invariant metric*. This metric is highly symmetric and unique up to a scalar multiple.

Note: The hight degree of symmetry is clear from the above discussion, but we haven't shown uniqueness.

8.7 Kinetic Energy and Geodesic Motion

For a particle moving on G along the trajectory $g(t)$, we can define the kinetic energy:

$$T = -\operatorname{Tr}\left(\dot{g}g^{-1}\dot{g}g^{-1}\right), \text{ with } \dot{g} \equiv dg/dt$$

and the action:

$$S = -\int_{t_1}^{t_0} \operatorname{Tr}\left(\dot{g}g^{-1}\dot{g}g^{-1}\right).$$

Exercise: (Example Sheet 3, Problem 9) Show that $\delta S \stackrel{!}{=} 0$ gives $\frac{d}{dt}(g^{-1}\dot{g}) = 0$ as the equation of motion using the above action. Solutions will correspond to motion along geodesics at constant speed.

If we integrate the equation of motion once with respect to time, we see:

$$\int \frac{d}{dt}(g^{-1}\dot{g})dt \implies g^{-1}\dot{g} = X_0 = \text{const} \in L(G).$$

This is another differential equation, which we can solve with the following function: $g(t) = \exp(tX_0)$ with $g(0) = I, \dot{g}(t)|_{t=0} = X_0$ (arbitrary initial velocity). It is easy to see that this function solves the differential equation if one recalls that $\exp(\pm tX_0)$ and X_0 commute. However, this special solution is not the most general solution, which is given by the left translation of the special solution $g(t) = g_0 \exp(tX_0)$, where $X_0 \in L(G), g_0 = \text{const} \in G$. We can check that this does in fact solve the differential equation:

$$\begin{aligned}
\dot{g}(t) &= g_0 X_0 \exp(tX_0) \\
g^{-1}\dot{g} &= (g_0 \exp(tX_0))^{-1} g_0 X_0 \exp(tX_0) \\
&= \exp(-tX_0) g_0^{-1} g_0 X_0 \exp(tX_0) \\
&= X_0, \text{ which was to be shown.}
\end{aligned}$$

Remark: Geodesic motion possesses enough symmetry that it is completely integrable.

8.8 $SU(2)$ Metric and Volume Form

Recall (from example sheets) the following parametrizations of $SU(2)$:

$$U = \begin{pmatrix} a_0 + ia_3 & ia_1 + a_2 \\ ia_1 - a_2 & a_0 - ia_3 \end{pmatrix}, \text{ with } a_0^2 + \vec{a}\cdot\vec{a} = 0,\ a_i \in \mathbb{R}$$
$$U = \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix}, \text{ with } |\alpha|^2 + |\beta|^2 = 1.$$

(These parametrizations are really the same and in fact give a quick way to recall explicit forms for the Pauli matrices.) We know that a $SU(2)\times SU(2)$ invariant metric is $SO(4)$ invariant. This fact follows because we showed that $\mathfrak{so}(4) = \mathfrak{su}(2)\oplus\mathfrak{su}(2)$, and the metric is a function $L(G)\times L(G) \longrightarrow \mathbb{R}$ or $\mathbb{C}$. We can consider the round metric on S^3:

$$ds^2 = da_0^2 + d\vec{a}^2, \text{ restricted to } S^3$$
$$ds^2 = d\alpha d\alpha^* + d\beta d\beta^*, \text{ with the restriction } |\alpha|^2 + |\beta|^2 = 1$$

Exercise: (Example Sheet 4, Problem 2) Show that $ds^2 = -\frac{1}{2}\operatorname{Tr}\left((dU)U^{-1}(dU)U^{-1}\right)$.

8.8.1 Euler Angle Parametrization

Let $\alpha = \cos\frac{\theta}{2}\exp\left(\frac{i}{2}(\phi+\psi)\right)$, $\beta = \sin\frac{\theta}{2}\exp\left(\frac{i}{2}(-\phi+\psi)\right)$, with $\theta\in[0,\pi]$, $\phi\in[0,2\pi]$, $\psi\in[0,4\pi]$. If one computes the differentials $d\alpha$ and $d\beta$, one finds after about six lines of easy algebra that:

$$ds^2 = d\alpha d\alpha^* + d\beta d\beta^* = \frac{1}{4}(d\theta^2 + d\phi^2 + d\psi^2 + 2\cos\theta\, d\phi d\psi).$$

This is the round metric on S^3 via Euler angles. (cf. Landau and Lifschitz Vol.I for more discussion.) The metric tensor is then:

$$g_{\mu\nu} = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \cos\theta \\ 0 & \cos\theta & 1 \end{pmatrix}.$$

The volume element is given by: $dV = \sqrt{\det g_{\mu\nu}}\, d\theta d\phi d\psi = \frac{1}{8}\sin\theta\, d\theta d\phi d\psi$. This notion of a volume form is useful in $SU(2)$ gauge theory quantization. With it, we can compute the total volume:

$$V = \frac{1}{8}\int \sin\theta\, d\theta d\phi d\psi = 2\pi^2$$

(Note: In general a volume form on a differentiable manifold $\mathcal{M}$ is nowhere-vanishing differential form of top degree. On a manifold $\mathcal{M}$ of $\dim n$, a volume form is an n-form, a section of the line bundle $\Omega^n(\mathcal{M}) = \Lambda^n(T^*\mathcal{M})$.)

Chapter 9

$SU(3)$ and its Representations

9.1 Roots

Recall the $SU(2)$ Lie algebra $\mathfrak{su}(2)$:

$$\begin{aligned}
[h, e_+] &= e_+ \\
[h, e_-] &= -e_- \\
[e_+, e_-] &= 2h,
\end{aligned}$$

where $\{e_\pm, h\}$ is a basis for $\mathfrak{su}(2)$. Recall that $\mathfrak{su}(3)$ consists of traceless, anti-hermitian 3×3 matrices. A convenient basis (and one with similarities to the basis for $\mathfrak{su}(2)$ above) for the complexified $\mathfrak{su}(3)$ (i.e., for $\mathfrak{su}(3) \otimes_{\mathbb{R}} \mathbb{C}$) is:

$$e_\alpha = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, e_{-\alpha} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, e_\beta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

$$e_{-\beta} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, e_\gamma = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, e_{-\gamma} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

$$h_\alpha = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, h_\beta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & -1/2 \end{pmatrix}, h_\gamma = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1/2 \end{pmatrix}.$$

With these definitions, we have the following commutation relations:

$$\begin{aligned}
[h_\alpha, e_\alpha] &= e_\alpha \\
[h_\alpha, e_{-\alpha}] &= -e_{-\alpha} \\
[e_\alpha, e_{-\alpha}] &= 2h_\alpha.
\end{aligned}$$

These are precisely the commutation relations fulfilled by $\mathfrak{su}(2)$. Thus $\{e_\alpha, e_{-\alpha}, h_\alpha\}$ form an $\mathfrak{su}(2)$ subalgebra, and similarly for β, γ. However, $h_\alpha + h_\beta = h_\gamma$, so these subalgebras are not linearly independent. We can choose an orthonormal basis for the diagonal matrices:

$$h_1 = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, h_2 = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

They are orthonormal in the sense: $\mathrm{Tr}({h_1}^2) = \mathrm{Tr}({h_2}^2) = 1/2$ and $\mathrm{Tr}(h_1 h_2) = 0$. We write $\vec{h} = (h_1, h_2)$ as a vector and define:

$$\begin{aligned}
\vec{\alpha} &= (1, 0) \\
\vec{\beta} &= (-1/2, \sqrt{3}/2) \\
\vec{\gamma} &= (1/2, \sqrt{3}/2).
\end{aligned}$$

Using this notation we see:

$$\begin{aligned}
h_\alpha &= \vec{\alpha} \cdot \vec{h} = (1, 0)(h_1, h_2) = h_1 \\
h_\beta &= \vec{\beta} \cdot \vec{h} = (-1/2, \sqrt{3}/2)(h_1, h_2) = -1/2h_1 + \sqrt{3}/2h_2 \\
&= \begin{pmatrix} -1/4 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & -1/2 \end{pmatrix} = h_\beta \\
h_\gamma &= \vec{\gamma} \cdot \vec{h} = (1/2, \sqrt{3}/2)(h_1, h_2) = -1/2h_1 + \sqrt{3}/2h_2 \\
&= \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & -1/4 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & -1/2 \end{pmatrix} = h_\gamma .
\end{aligned}$$

More compactly, we can write:

$$\begin{aligned}
\mathrm{ad}_{\vec{h}}\, e_\alpha &= (1, 0) = \vec{\alpha} \\
\mathrm{ad}_{\vec{h}}\, e_\beta &= \left(\frac{-1}{2}, \frac{\sqrt{3}}{2} \right) = \vec{\beta} \\
\mathrm{ad}_{\vec{h}}\, e_\gamma &= \left(\frac{+1}{2}, \frac{\sqrt{3}}{2} \right) = \vec{\gamma}
\end{aligned}$$

In other words, all of this notation is consistent, which is always a comforting place to begin. This notation leads to the *root diagram* of $\mathfrak{su}(3)$:

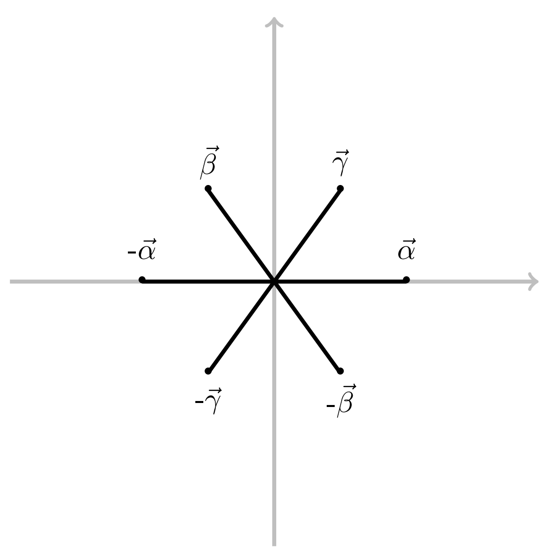


Figure 9.1: The root diagram of $L(SU(3))$

$\pm\vec{\alpha}, \pm\vec{\beta}, \pm\vec{\gamma}$ are the *roots* of $\mathfrak{su}(3)$. (In this case the are of unit length.) The negative roots are included because:

$$[e_\alpha, e_{-\alpha}] = -2h_\alpha = 2h_{-\alpha} = 2(-\vec{\alpha}\cdot\vec{h}),$$

and so on. $\mathfrak{su}(3)$ is said to have *rank* 2. The rank is the dimension of the maximum commuting subalgebra (the Cartan subalgebra, or CSA), which is here spanned by h_1 and h_2. The root diagram is two-dimensional. $\mathfrak{su}(3)$ has three (special) $\mathfrak{su}(2)$ subalgebras, one for each pair $\pm\alpha, \pm\beta, \pm\gamma$. There are also further brackets that we have not given, e.g.,

$$\begin{aligned}
&[e_\alpha, e_\alpha] = 0,\\
&[e_\alpha, e_\beta] = e_\gamma,\\
&[e_\alpha, e_\gamma] = 0,\\
&\text{and so on}\ldots
\end{aligned}$$

Remark. This discussion of roots and the associated theory is not complete due to the time constrains of the lecture.

Remark. The roots above are consistent with our previous definition of roots as generalized eigenvalues of the Cartan subalgebra in the adjoint representation.

9.2 Representations and Weights

Suppose d is a representation of $\mathfrak{su}(3)$ acting on V. Let h_1 and h_2 be as before. Then $d(h_1)$ and $d(h_2)$ commute (essentially be definition, as they span the Cartan subalgebra) and can be simultaneously diagonalized. Thus V decomposes into subspaces. Let

$$V_{\lambda_1,\lambda_2} = \{v \in V : d(h_i)v = \lambda_i v\}$$

(i.e., V_{λ_1,λ_2} is the span of the simultaneous eigenvectors belonging to eigenvalues λ_1 and λ_2 of $d(h_1)$ and $d(h_2)$. We can simplify notation by writing $\underline{h} = (h_1, h_2), \underline{\lambda} = (\lambda_1, \lambda_2)$. Then

$$V_{\underline{\lambda}} = \{v \in V : d(\underline{h})v = \underline{\lambda}v\}$$

If $V_{\underline{\lambda}}$ is a non-zero subspace, we say that $\underline{\lambda}$ *is weight.* In this case, $V_{\underline{\lambda}}$ *is a weight space for* d.

Example: The fundamental representation of $\mathfrak{su}(3)$.
The fundamental representation of $\mathfrak{su}(3)$ is 3-dimensional, and

$$h_1 = \begin{pmatrix} 1/2 & 0 & 0\\ 0 & -1/2 & 0\\ 0 & 0 & 0\end{pmatrix}, h_2 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1/2 & 0 & 0\\ 0 & 1/2 & 0\\ 0 & 0 & -1\end{pmatrix}$$

are diagonal. Thus we can immediately read off the eigenvalues and hence the weights:

$$\begin{aligned}
&\left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right),\\
&\left(-\frac{1}{2}, \frac{1}{2\sqrt{3}}\right),\\
&\left(0, -\frac{1}{\sqrt{3}}\right).
\end{aligned}$$

The weight spaces are clearly one-dimensional and given by:

$$\begin{pmatrix} c \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ c \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}.$$

We can show the weights on the root diagram:

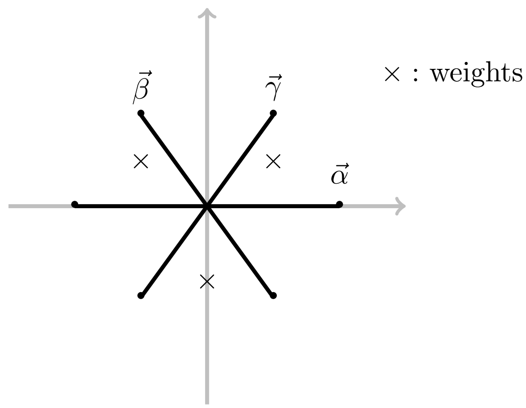


Figure 9.2: The root diagram of $L(SU(3))$ along with the weights

Note: The weights form an equilateral triangle
Note: We have 3 weights since the fundamental representation is itself 3-dimensional. Each weight is 2-dimensional, since the Cartan subalgebra is 2-dimensional.

9.2.1 General Constraint on Weights

We saw previously that $\{e_\alpha, e_{-\alpha}, h_\alpha\}$ span a $\mathfrak{su}(2)$ subalgebra of $\mathfrak{su}(3)$. Using restriction, we see that the representation d gives a representation of $\mathfrak{su}(2)$ acting on V. From our experience with $SU(2)$ theory (think angular momentum), we know that $d(h_\alpha)$ has half-integer or integer eigenvalues. If $v \in V_{\underline{\lambda}}$, then

$$\begin{aligned} d(h_\alpha)v &= d(\underline{\alpha} \cdot \underline{h})v, \text{ def of } h_\alpha \\ &= \underline{\alpha} \cdot d(\underline{h})v, \text{ since } \underline{\alpha} \text{ are numbers} \\ &= \underline{\alpha} \cdot \underline{\lambda} v, \text{ since } v \in V_{\underline{\lambda}} \end{aligned}$$

This result combined with our insight from $SU(2)$ theory tells us:

$$2\underline{\alpha} \cdot \underline{\lambda} \in \mathbb{Z} \tag{9.1}$$

The same argument and result holds for any root $\underline{\delta}$ (i.e., $\pm\underline{\alpha}, \pm\underline{\beta}, \pm\underline{\gamma}$) of $\mathfrak{su}(3)$. Thus the weights λ, for *any* representation d of $\mathfrak{su}(3)$ are constrained by (9.1). The possible weights form a lattice. The basis lattice vectors are $\underline{w}_1 = \left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right)$ and $\underline{w}_2 = \left(0, \frac{1}{\sqrt{3}}\right)$, which satisfy:

$$\begin{aligned} 2\underline{\alpha} \cdot \underline{w}_1 &= 1 \\ 2\underline{\alpha} \cdot \underline{w}_2 &= 0 \\ 2\underline{\beta} \cdot \underline{w}_1 &= 0 \\ 2\underline{\beta} \cdot \underline{w}_2 &= 1 \end{aligned}$$

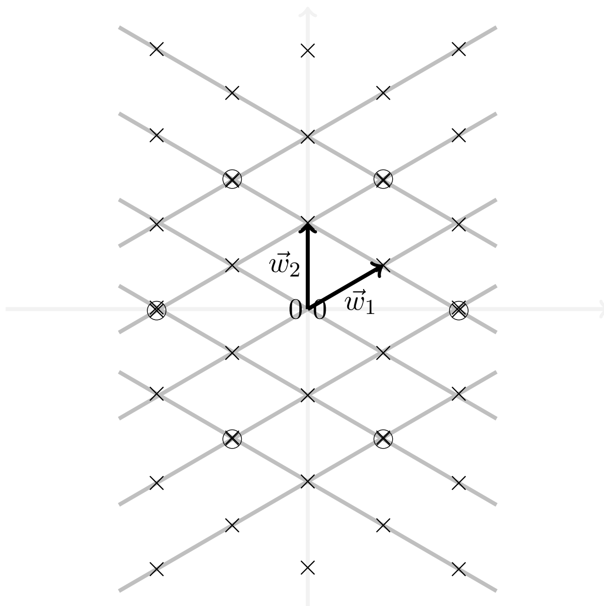


Figure 9.3: The weight lattice

(Prove by looking at the definitions of $\underline{\alpha}, \underline{\beta}$ previously.) A general weight is $\underline{\lambda} = n_1\underline{w}_1 + n_2\underline{w}_2$, where $n_1, n_2 \in \mathbb{Z}$. This satisfies the constraint equation (9.1) for $\underline{\delta} = \underline{\alpha}$ or $\underline{\delta} = \underline{\beta}$ and for $\underline{\gamma} = \underline{\alpha} + \underline{\beta}$. This is the *weight lattice*. The weights of $d = \text{ad}$ (the adjoint representation) are the roots and 0. They are shown by the circles in the lattice above.

Note: Roots are weights, since ad is of course a representation. In the case of $\mathfrak{su}(3)$,

$$\begin{aligned}
&\text{ad}_{h_1} e_{\pm\alpha} = [h_1, e_{\pm\alpha}] = \pm e_{\pm\alpha}\\
&\text{ad}_{h_2} e_{\pm\alpha} = [h_2, e_{\pm\alpha}] = 0\\
&\implies \text{ad}_{\underline{h}}\, e_{\pm\alpha} = (\pm 1, 0)e_{\pm\alpha}, \text{ which just says that the weight is } \pm\underline{\alpha}.
\end{aligned}$$

Similarly, we see the other weights are determined according to:

$$\begin{aligned}
\text{ad}_{\underline{h}}\, e_{\pm\beta} &= (\mp 1/2, \pm\sqrt{3}/2)e_{\pm\beta} = \pm\underline{\beta} e_{\pm\beta}\\
\text{ad}_{\underline{h}}\, e_{\pm\gamma} &= (\pm 1/2, \pm\sqrt{3}/2)e_{\pm\gamma} = \pm\underline{\gamma} e_{\pm\gamma}\\
\text{ad}_{\underline{h}}\, h_1 &= 0\\
\text{ad}_{\underline{h}}\, h_2 &= 0
\end{aligned}$$

Thus ad has 8 weights $\{\pm\underline{\alpha}, \pm\underline{\beta}, \pm\underline{\gamma}, \underline{0}, \underline{0}\}$.

9.2.2 Weights of Representations of $\mathfrak{su}(3)$

Weights belong to the weight lattice. Thus must also form *complete strings* of weighs for each $\mathfrak{su}(2)$ subalgebra. Some possibilities (irreps of $\mathfrak{su}(3)$) are given below.

9.2.3 Conjugate Representations

The fundamental representation of $SU(3)$ is $D(U) = U$. The conjugate representation is $D(U) = U^*$. This defines a representation because $U_1U_2 = U_3 \mapsto U_1^*U_2^* = U_3^*$. We may also consider the

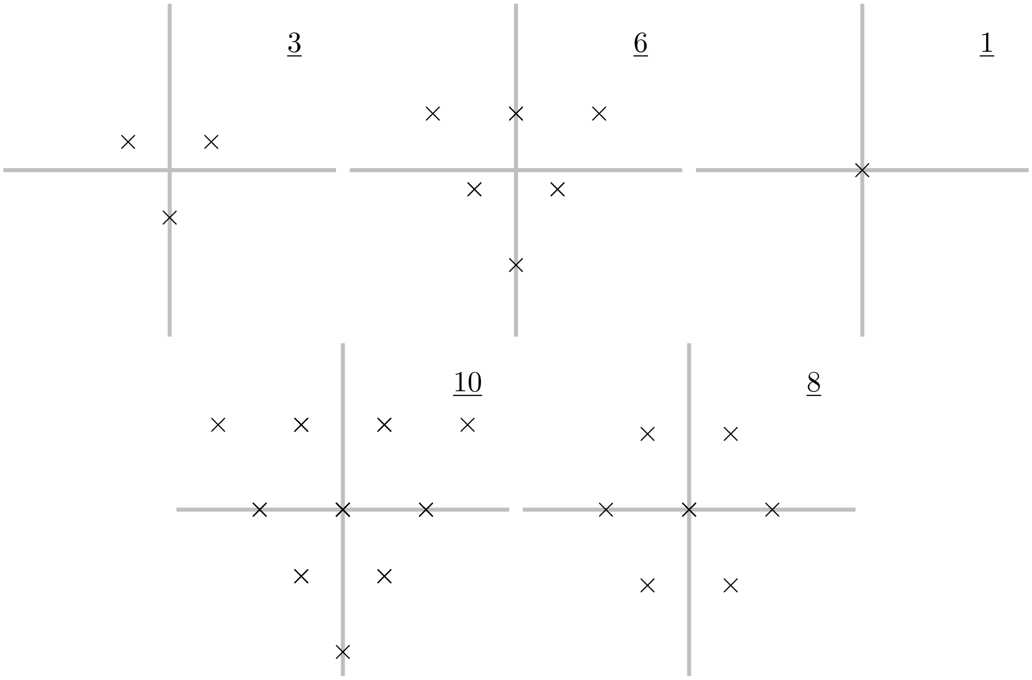


Figure 9.4: The weight diagrams for some irreducible representations of $L(SU(2))$

associated representations of the Lie algebra:

Fundamental:	$D(I+X) = I + X$, so $d(X) = X$
Cojugate:	$D(I+X) = I + X^*$, so $d(X) = X^*$

Since $X \in \mathfrak{su}(3)$, it is of course anti-hermitian and thus has imaginary eigenvalues: $Xv = i\mu v, \mu \in \mathbb{R}$. Then for the conjugate: $X^* v^* = -i\mu v^*$, and so we see that the eigenvalues change sign. Now we move to the complexification $\mathfrak{su}(3) \otimes_{\mathbb{R}} \mathbb{C}$, i.e., we now allow for complex linear combination such as e_α, h_α, and so on. The weight diagrams of $\underline{3}$ and $\overline{\underline{3}}$ are:

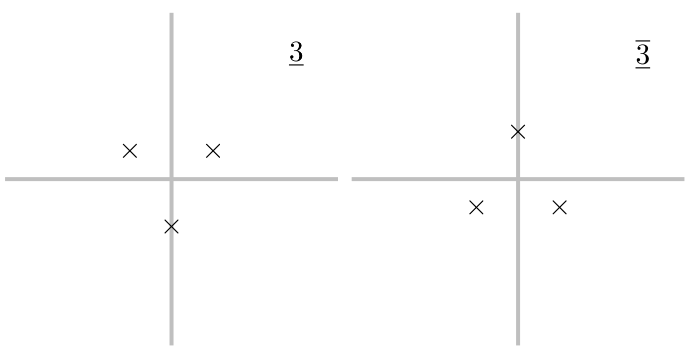


Figure 9.5: The weight diagrams of the fundamental and conjugate representations of $L(SU(3))$

Note that these are different, so $\underline{3}$ and $\overline{\underline{3}}$ are *not* the same. Compare this to the fundamental representation of $\mathfrak{su}(2)$, which is self-conjugate.
Similarly, the conjugates of $\underline{6}$ and $\underline{10}$ are distinct, $\overline{\underline{6}}$ and $\overline{\underline{10}}$, respectively. However, $\underline{8} = \overline{\underline{8}}$ is self-conjugate.

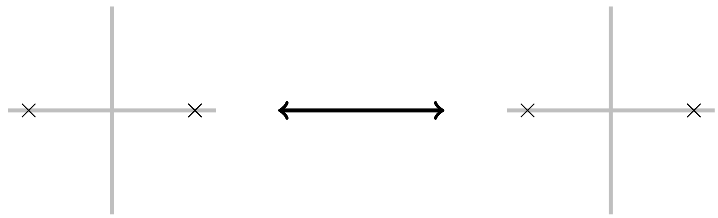
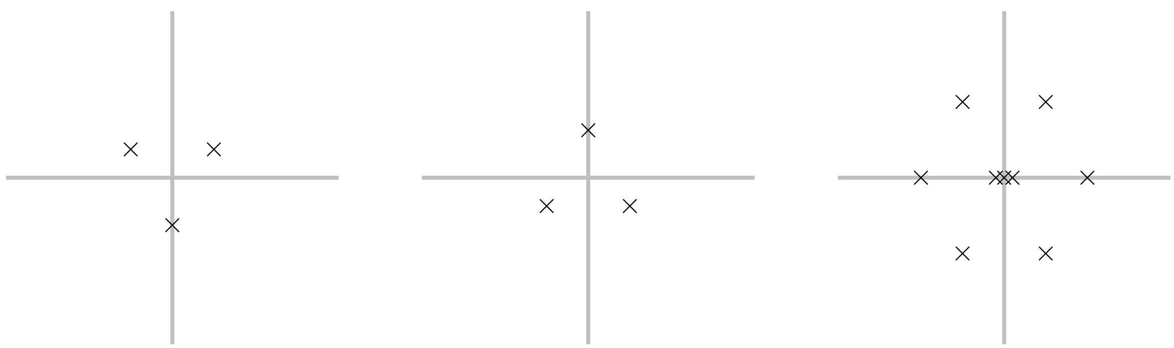


Figure 9.6: The weight diagrams for the fundamental and conjugate representations of $L(SU(2))$.

9.2.4 Tensor Products for $\mathfrak{su}(3)$

We can construct all irreducible representations of $\mathfrak{su}(3)$ from tensor products of $\underline{3}$ (the fundamental representation) and $\overline{\underline{3}}$ (its conjugate) and then reducing the tensor product to irreducible representations. One *adds weights* to identify the representation present in the tensor product.

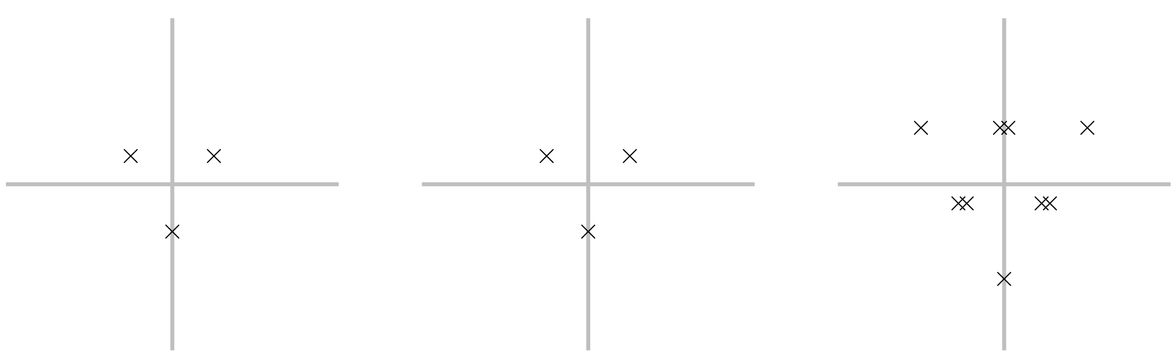
Example 1: $\underline{3} \otimes \overline{\underline{3}} = \underline{8} \oplus \underline{1}$



$$\underline{3} \qquad \otimes \qquad \overline{\underline{3}} \qquad = \qquad \underline{8} \oplus \underline{1}$$

The singlet is invariant under the action of $SU(3)$. If a tensor has the form $v^*_\beta u_\alpha$, then the singlet is $\sum_{\alpha=1}^{3} v^*_\alpha u_\alpha = v^\dagger u$.

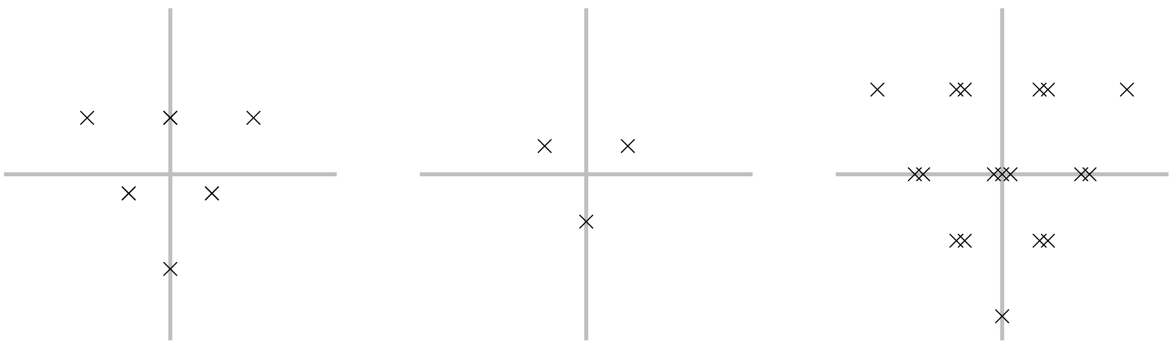
Example 2: $\underline{3} \otimes \underline{3} = \underline{6} \oplus \overline{\underline{3}}$



$$\underline{3} \qquad \otimes \qquad \underline{3} \qquad = \qquad \underline{6} \oplus \overline{\underline{3}}$$

$\underline{6}$ is the symmetric 3×3 tensor, $S_{\alpha\beta}$, and $\overline{\underline{3}}$ is the antisymmetric 3×3 tensor, $A_{\alpha\beta}$. Since these are acted upon by $SU(3)$, these are complex.

Example 3: $\underline{3}\otimes\underline{3}\otimes\underline{3} = (\underline{6}\oplus\overline{\underline{3}})\otimes\underline{3} = (\underline{6}\otimes\underline{3})\oplus(\underline{8}\otimes\underline{1}) = \underline{10}\oplus\underline{8}\oplus\underline{8}\oplus\underline{1}$



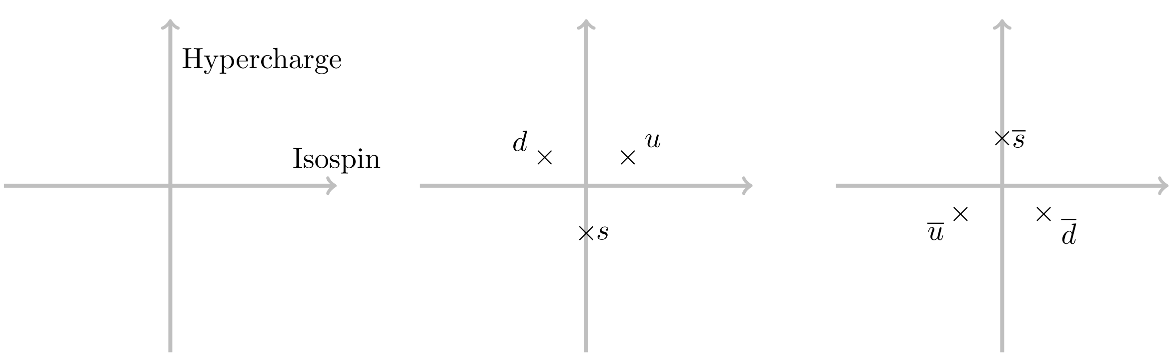
$$\underline{6} \qquad \otimes \qquad \underline{3} \qquad = \qquad \underline{10}\oplus\underline{8}$$

$\underline{10}$ is a totally symmetric tensor. $\underline{1}$ is a totally antisymmetric tensor. The singlet is $\epsilon_{\alpha\beta\gamma}u^\alpha v^\beta w^\gamma$.

9.3 Quarks

Heisenberg noted close degeneracy (in the physical sense) of the proton and neutron masses $(p,n)^T = (938\text{ MeV}, 940\text{ MeV})^T$ and observed that (with the exception of charge) the two particles had similar properties in nuclei. This lead him to propose *isospin symmetry*, an $SU(2)$ symmetry with $(p,n)^T$ as a doublet. Isospin was confirmed by the discovery of 3 pions $(\pi^+,\pi^0,\pi^-)^T$ and interactions that were approximately $SU(2)$ invariant. The discovery of more particles, like $(K^+,K^-)^T$, and the new *strangeness* conservation led Gell-Mann to suggest a larger $SU(3)_{\text{flavor}}$ symmetry with $SU(2)_{\text{isospin}}$ as a subgroup. The particles transforming under the *fundamental representation* ($\underline{3}$) of $SU(3)$ are now understood to be the quark $(u,d,s)^T$, i.e., $(1,0,0)^T$ is identified with u quark states and so on. All other hadrons are multiquark state in $SU(3)$ multiplets. Mathematically, we use tensor products to produce these multiplets. Particles in irreducible representations have approximately the *same mass*. (Different irreducible representations are *not* related by $SU(3)$ symmetry.)

In the multiplets, quarks q are in $\underline{3}$, while antiquarks $\bar{q}$ are in $\overline{\underline{3}}$



The masses of the up and down quarks $\begin{pmatrix} u \\ d \end{pmatrix}$ are very similar (and small, on the order of a few

MeV). The strange quark (s) has a "large" mass (on the order of 150 MeV) (Remark: The particle data group says 95 ± 5 MeV for the mass of the strange quark.)

9.3.1 Meson Octet

The meson octet comes from the tensor product $\underline{3} \otimes \overline{\underline{3}}$.

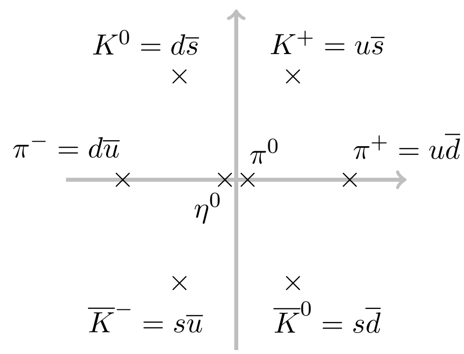


Figure 9.7: The meson octet

This agrees with what is seen: 4 kaons (mass ~500 MeV), 3 pions (mass ~140 MeV), 1 eta (mass ~500 MeV). There is also a further *singlet* η' (also neutral, but heavier than η_0).

9.3.2 Baryon Octet and Decuplet

We obtain $\underline{8}$ and $\underline{10}$ from $\underline{3} \otimes \underline{3} \otimes \underline{3}$ (qqq).

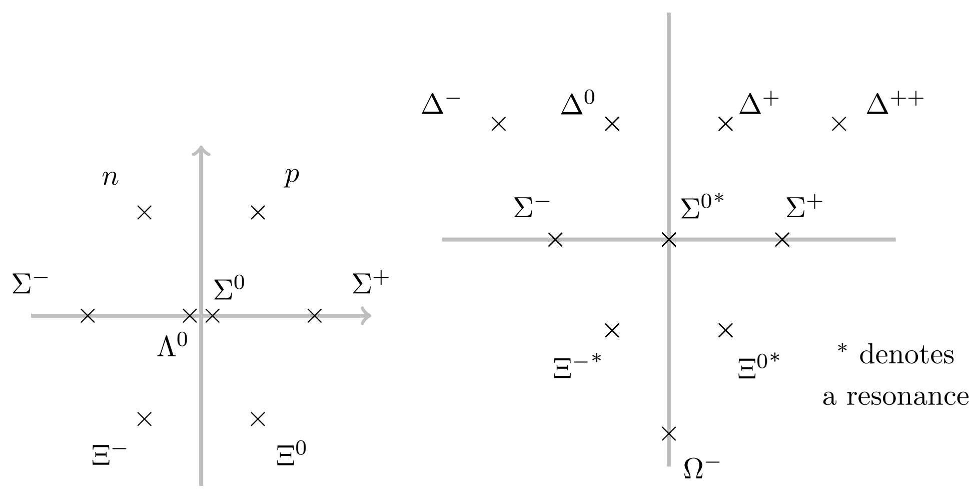
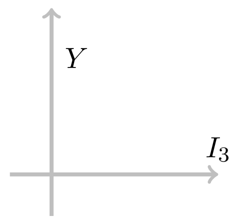


Figure 9.8: The baryon octet and decuplet

The diagrams above are consistent with what is seen at hadron colliders. The baryon octet contains the long-lived baryons, which have spin 1/2 and decay weakly. The baryon decuplet consists mostly of resonances. They are short-lived and have spin 3/2. The Ω^- was a successful prediction of this scheme and is the only long-lived one in this group.

There are 3 conserved quantities in *strong interactions*:

- Y (related to strangeness), hypercharge
- I_3 (related to electric charge), 3rd component of isospin
- B (Baryon number)



I_3 and Y are the eigenvalues of $d(h_1)$ and $2/\sqrt{3}d(h_2)$, respectively. The net quark numbers

$$\begin{aligned} N_u &= \#u - \#\bar{u} \\ N_d &= \#d - \#\bar{d} \\ u_s &= \#s - \#\bar{s} \end{aligned}$$

are all conserved under strong interactions. We can also have the following relationships:

$$\begin{aligned} &B \equiv 1/3(N_u + N_d + N_3) \\ &I_3 = 1/2N_u - 1/2N_d \\ &Y = 1/3N_u + 1/3N_d - 2/3N_s \\ &Q = 2/3N_u - 1/3N_d - 1/3N_s \\ &S = -N_s. \end{aligned}$$

Of these five related quantum numbers, only three are independent. One relation is $Q = I_3 + 1/2Y$. For the quarks, $Q(U) = 2/3$, $Q(d) = -1/3$, $Q = -1/3$. Originally thought to be algebraic curiosities, these values have found verification from deep inelastic scattering experiments.

9.3.3 The Pauli Principle and Color

Consider $\Delta^{++} = u^{\uparrow}u^{\uparrow}u^{\uparrow}$, consisting of 3 spin-up quarks. This is $|3/2, 3/2\rangle$ state. Through models, we believe that this state should have a spatially symmetric wavefunction. Since it also has a symmetric spin and flavor state, this appears to violate the Pauli principle for spin-1/2 quarks. This lead us to propose a further "color" label for quarks, $q_\nu, \nu = 1, 2, 3$. States such as qqq must be totally antisymmetric in color: $\epsilon_{\mu\nu\sigma}q_\mu q_\nu q_\sigma$. This leads to the $SU(3)$ gauge theory, QCD, where color symmetry rules become a consequence of gauge invariance. This also explains why only $\underline{3} \otimes \overline{\underline{3}}$ $(q_\nu \bar{q}_\nu)$, $\underline{3} \otimes \underline{3} \otimes \underline{3}$ (qqq), and $\overline{\underline{3}} \otimes \overline{\underline{3}} \otimes \overline{\underline{3}}$ $(\bar{q}\bar{q}\bar{q})$are seen physically. Note that free quaks q_ν are not gauge invariant. Rather surprisingly, there is no evidence for $qq\bar{q}\bar{q}$ or $qqqqqq$ states, and there are seemingly no glueballs. "The dynamics of color confinement remain mysterious."

Chapter 10

Complexification of $L(G)$, Representations

10.1 $L(G)^{\mathbb{C}}$

Consider a real Lie algebra $L(G)$ with basis $\{T_k\}$ and brackets $[T_i, T_j] = c_{ijk}T_k$, where $c_{ijk} \in \mathbb{R}$. Here the T's themselves need *not* be real matrices, but the structure constants *are* real. The full Lie algebra then consists of the real Linear span of the basis matrices. The *complexification* of $L(G)$, denoted $L(G)^{\mathbb{C}}$, has a general element $\sum_k \lambda_k T_k$, with $(\lambda_k \in \mathbb{C})$. The bracket is unchanged:

$$[\lambda_i T_i, \mu_j T_j] = \lambda_i \mu_j c_{ijk} T_k.$$

A representation d of $L(G)$ becomes a rep d of $L(G)^{\mathbb{C}}$:

$$d(\lambda_k T_k) = \lambda_k d(T_k)$$

Here *the real dimension doubles*, while the *complex dimension is the same* as the original Lie algebra.

10.2 $L(G)^{\mathbb{C}}$ as a real Lie algebra $\Re\{L(G)^{\mathbb{C}}\}$

As a real algebra, $\dim(L(G)^{\mathbb{C}}) = 2\dim(L(G))$. A basis for $L(G)^{\mathbb{C}}$ is given by $\{X_k = T_k, Y_k = iT_k\}$. (We assume here that these are all independent, i.e., that multiplication by i does not reproduce a basis element.) We have brackets:

$$\begin{aligned}
[X_i, X_j] &= c_{ijk} X_k \\
[X_i, Y_j] &= c_{ijk} Y_k \\
[Y_i, Y_j] &= -c_{ijk} X_k
\end{aligned}$$

We can produce three types of representations of $L(G)^{\mathbb{C}}$ using the representation d of $L(G)$:

(i) $d(X_k) = d(T_k)$
$d(Y_k) = id(T_k)$.
This is the representation of $L(G)^{\mathbb{C}}$ we considered above

(ii) $d(X_k) = d(T_k)$
$d(Y_k) = -id(T_k)$.
This representation is conjugate to (i).

Note that these representation preserve the bracket, as they must. We can confirm this:

(a)

$$\begin{aligned}[d(X_i), d(Y_j)] &= [d(T_i), id(T_j)] \\ &= d([T_i, iT_j]) = d(ic_{ijk}T_k) \\ &= c_{ijk}d(Y_k) \\ [d(Y_i), d(Y_j)] &= [id(T_i), id(T_j)] \\ &= d([iT_i, iT_j]) = d(-c_{ijk}T_k) \\ &= -c_{ijk}d(X_k)\end{aligned}$$

(b)

$$\begin{aligned}[d(X_i), d(Y_j)] &= [d(T_i), -id(T_j)] \\ &= d([T_i, -iT_j]) = d(-ic_{ijk}T_k) \\ &= c_{ijk}d(Y_k) \\ [d(Y_i), d(Y_j)] &= [-id(T_i), -id(T_j)] \\ &= d([-iT_i, -iT_j]) = d(-c_{ijk}T_k) \\ &= -c_{ijk}d(X_k)\end{aligned}$$

We can combine the first two representation via a tensor product to get a third representation. Start with $d^{(1)}, d^{(2)}$ as representations of $L(G)$:

(iii) Consider the representation of $L(G)^{\mathbb{C}} = \text{span}\{X_k, Y_k\}$ defined by:

$$\begin{aligned}d(X_k) &= d^{(1)}(T_k) \otimes I + I \otimes d^{(2)}(T_k) \\ d(Y_k) &= i(d^{(1)}(T_k) \otimes I - I \otimes d^{(2)}(T_k))\end{aligned}$$

In this way we can construct an irreducible representation d of $\Re\{L(G)^{\mathbb{C}}\}$ from the irreducible representations $d^{(1)}, d^{(2)}$ of $L(G)$.

10.3 Another Point of View

From this algebra we can construct two commuting copies of the $L(G)$ algebra:

$$\begin{aligned}Z_k &= \frac{1}{2}(X_k - iY_k) \\ \tilde{Z}_k &= \frac{1}{2}(X_k + iY_k)\end{aligned}$$

One easily computes the brackets and sees:

$$\begin{aligned}[Z_i, Z_j] = Z_iZ_j - Z_jZ_i &= \frac{1}{4}((X_i - iY_i)(X_j - iY_j) - (X_j - iY_j)(X_i - iY_i)) \\ &= \frac{1}{4}(X_iX_j - iX_iY_j - iY_iX_j - Y_iY_j - X_jX_i + iX_jY_i + iY_jX_i + Y_jY_i) \\ &= \frac{1}{4}([X_i, X_j] - i[X_i, Y_j] + i[X_j, Y_i] - [Y_j, Y_i]) \\ &= \frac{1}{4}(c_{ijk}X_k - ic_{ijk}Y_k + ic_{jik}Y_k + c_{ijk}X_k) \\ &= \frac{1}{2}c_{ijk}(X_k - iY_k) \\ &= c_{ijk}Z_k\end{aligned}$$

By a similar calculation we see $[\tilde{Z}_i, \tilde{Z}_j] = c_{ijk}\tilde{Z}_k$ and $[Z_i, \tilde{Z}_j] = 0$. We represent $\{Z_k\}$ using $d^{(1)}$ and $\{\tilde{Z}_k\}$ using $d^{(2)}$ to get the tensor products:

$$\begin{aligned} d(Z_k) &= d^{(1)}(T_k) \otimes I \\ d(\tilde{Z}_k) &= I \otimes d^{(2)}(T_k). \end{aligned}$$

This then gives the same result as (*iii*). We have just used the complexification of $\Re\{L(G)^{\mathbb{C}}\}$. This is a bit complicated, but it is exactly what is need to understand the Lorentz group and its representation.

Example:

$$\begin{aligned} &\mathfrak{su}(n) = \{\text{traceless anti-hermitian matrices}\} \\ &\mathfrak{su}(n)^{\mathbb{C}} = \{\text{traceless complex matrices}\} \\ &\Re\{\mathfrak{su}(n)^{\mathbb{C}}\} = \{\text{traceless anti-hermitian matrices, traceless hermitian matrices}\} \end{aligned}$$

Remark: Previously in our discussion of $\mathfrak{su}(3)$ we had said that the 3×3 matrices $\{e_{\pm\alpha}, e_{\pm\beta}, e_{\pm\gamma}, h_\alpha, h_\beta, h_\gamma\}$ were a basis for $\mathfrak{su}(n)^{\mathbb{C}}$, which is consistent with the example above.

Chapter 11

Lorentz Group and Lie Algebra, Representation

Lorentz transformations act on 4-vectors $x^\mu \to L^\mu{}_\nu x^\nu, \mu, \nu = 0, 1, 2, 3$. Lorentz transformations by definition preserve the square of the relativistic interval $s^2 \equiv \eta_{\mu\nu} X^\mu X^\nu$, where $\eta_{\mu\nu} = \eta^{\mu\nu} = diag(1, -1, -1, -1)$. The metric signature is (1,3). Preservation of the relativist interval is the condition $s^2 \to s'^2 \stackrel{!}{=} s^2$, which implies:

$$L^\mu{}_\sigma L^\nu{}_\tau \eta_{\mu\nu} = \eta_{\mu\nu}$$

This is the defining equation for the Lorentz group $O(1, 3)$. One notes that this group is "sort of a variant of $O(4)$." Near the group identity, $L^\mu{}_\sigma = \delta^\mu{}_\sigma + \epsilon\, l^\mu{}_\sigma$, where ϵ is infinitesimal. Substituting this into the fundamental equation, one finds after two lines of algebra:

$$l^\mu{}_\sigma \eta_{\mu\tau} + \eta_{\sigma\nu} l^\nu{}_\tau = 0.$$

Lowering indices, we see that $l_{\tau\sigma} + l_{\sigma\tau} = 0$, i.e., l is antisymmetric. Thus we may write

$$l^\mu{}_\sigma = \begin{pmatrix} 0 & a & b & c \\ a & 0 & d & -e \\ b & -d & 0 & f \\ c & e & -f & 0 \end{pmatrix}.$$

One observes that the first column and the first row are symmetric. This comes from the fact that we have written the mixed tensor, which comes from raising the first index with the Minkowski metric. Letting $a, b, c, d, e, f \in \mathbb{R}$, these matrices form a 6-dimensional real Lie algebra of O(1,3). A basis is given by:

$$K_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, K_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, K_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$J_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, J_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, J_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The K's are generators for the boosts, while the J's are generators for the boosts. We can combine these generators into $M^{\mu\nu} = -M^{\nu\mu}$ with $M^{0k} = K_k$, $M^{ij} = \epsilon_{ijk}$. Explicitly, we have:

$$\begin{pmatrix} 0 & K_1 & K_2 & K_3 \\ -K_1 & 0 & J_3 & -J_2 \\ -K_2 & -J_3 & 0 & J_1 \\ -K_3 & J_2 & -J_1 & 0 \end{pmatrix}$$

For J and K we have the following brackets:

$$[J_i, J_j] = \epsilon_{ijk} J_k$$
$$[J_i, K_j] = \epsilon_{ijk} K_k$$
$$[K_i, K_j] = -\epsilon_{ijk} J_k$$

The first of these is clear, since it's just the defining relationship for the Lie algebra $\mathfrak{so}(3)$. Together, the three relations form the Lorentz algebra. Note that the minus sign in the final line is critical. ("It's omission on Costa and Folgi is a scandalous error"). This is precisely the algebra of $\Re(\mathfrak{su}(2)^{\mathbb{C}})$ that we met previously. To construct an irreducible representation, use the spin j representation of $SU(2)$. Let $T_i, i = 1, 2, 3$ be the standard basis of $\mathfrak{su}(2)$:

$$d^{(j)}(J_i) = d^{(j)}(T_i)$$
$$d^{(j)}(K_i) = \pm i\, d^{(j)}(T_i).$$

Note that in the second line the plus sign corresponds to a type (i) representation of $L(G)^{\mathbb{C}}$ while the minus sign corresponds to a type (ii) representation.
A general irrep of the Lorentz algebra is a tensor product (j_1, j_2):

$$d^{(j_1,j_2)}(J_i) = d^{(j_1)}(T_i) \otimes I + I \otimes d^{(j_2)}(T_i)$$
$$d^{(j_1,j_2)}(K_i) = i\, d^{(j_1)}(T_i) \otimes I - i\, I \otimes d^{(j_2)}(T_i)$$

Remark: One can also consider things like $1/2(J_i + iK_i)$, but this lives in the complexification of $\Re(\mathfrak{su}(2)^{\mathbb{C}})$. This construction is admittedly a bit confusing, so we do our best to avoid it here.
The $\mathfrak{su}(2)$ subalgebra $\{J_i, i = 1, 2, 3\}$ is represented by the standard tensor product $j_1 \otimes j_2$.

Global Aspects
$O(1, 3)$ has four disconnected components. They are labelled by whether $\det L^\mu{}_\nu = \pm 1$ and by if $L^0{}_0 \geq +1$ or ≤ -1. G^{conn}, the part of O(1,3) connected to the identity, has $\det L = 1$ and $L^0{}_0 \geq 1$. This subgroup is called $SO(1, 3)^{\uparrow}$. The rotation subgroup in G^{conn} generated by $\{J_i\}$ is a copy of $SO(3)$. True representations of G^{conn} have integer spin. This requires $j_1 + j_2$ to be an integer (weights of $j_1 \otimes j_2, \dots$), e.g. $(j_1, j_2) = (0, 0)$ or $(1/2, 1/2)$. The Lorentz group has a double cover with spinor representations, e.g., (1/2,0).

Examples (of reps.)

(a) $(\frac{1}{2},0)$ spinor rep (spin $\frac{1}{2}$)
This corresponds to the fundamental representation of $\mathfrak{sl}(2, \mathbb{C})$

(b) $(0,\frac{1}{2})$ spinor rep (spin $\frac{1}{2}$)
This is conjugate to the fundamental representation of $\mathfrak{sl}(2, \mathbb{C})$

(c) $(\frac{1}{2},\frac{1}{2})$ is the 4-vector representation of the Lorentz group. Note that *under* $SO(3)$ *rotations*, $(\frac{1}{2},\frac{1}{2})$ is reducible: $\underline{\frac{1}{2}} \otimes \underline{\frac{1}{2}} = \underline{1} \oplus \underline{0}$. This corresponds to the 4-vector decomposition $(x^0, \vec{x})$. Note of course that $(\frac{1}{2},\frac{1}{2})$ is not reducible under general Lorentz transformations.

Note: (a) and (b) are left- and right-handed Weyl spinors. A Dirac spinor is $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$.

Chapter 12

Poincare Group and Particle States

The Poincare group (Poinc) combines Lorentz transformations and spacetime translations and acts transitively on Minkowski space $\mathcal{M}_4$. The isotropy group at one point (e.g., the origin) is the Lorentz group, so $\mathcal{M}_4 = Poinc/Lorentz$. (Let's think about this for a moment: Consider Minkowski space as an affine space instead of a vector space, since the full Poincare group involves translations. The familiar homogeneous Lorentz transformations are simply rotations - hyperbolic in the case of boosts - and so correspond to the usual linear transformations of Minkowski space as a vector space. Rotations clearly leave one point invariant, so we think of the Lorentz group as the isotropy group / stabilizer in this sense. The result is then clear.) More explicitly, elements of the Poincare group are $(L^\mu{}_\nu, a^\mu)$ and the action is $x^\mu \to L^\mu{}_\nu x^\nu + a^\mu$. There is a convenient 5×5 matrix realization of this:

$$\begin{pmatrix} x \\ 1 \end{pmatrix} \longrightarrow \begin{pmatrix} L & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} Lx + a \\ 1 \end{pmatrix},$$

where x and a are column 4-vectors and the above matrices are in block form. Using this notation group multiplication is given by:

$$\begin{pmatrix} L & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} L' & a' \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} LL' & La' + a \\ 0 & 1 \end{pmatrix}$$

Both the Lorentz group and the translation group are subgroups of the Poincare group, but they do not commute. The Lie algebra $L(Poinc)$ has the basis:

$$\tilde{M}^{\rho\sigma} = \begin{pmatrix} M^{\rho\sigma} & 0 \\ 0 & 0 \end{pmatrix}, \tilde{P}^\tau = \begin{pmatrix} 0 & P^\tau \\ 0 & 0 \end{pmatrix},$$

where $(M^{\rho\sigma})^\alpha{}_\beta = \eta^{\rho\alpha}\delta^\sigma{}_\beta - \eta^{\sigma\alpha}\delta^\rho{}_\beta$ and $(P^\tau)^\beta = \eta^{\tau\beta}$. The brackets for the Lie algebra are the following:

$$[\tilde{M}^{\rho\sigma}, \tilde{M}^{\tau\mu}] = \eta^{\sigma\tau}\tilde{M}^{\rho\mu} - \eta^{\rho\tau}\tilde{M}^{\sigma\mu} + \eta^{\rho\mu}\tilde{M}^{\sigma\tau} - \eta^{\sigma\mu}\tilde{M}^{\rho\tau} \tag{12.1}$$

$$[\tilde{M}^{\rho\sigma}, \tilde{P}^\tau] = \eta^{\sigma\tau}\tilde{P}^\rho - \eta^{\rho\tau}\tilde{P}^\sigma \tag{12.2}$$

$$[\tilde{P}^\tau, \tilde{P}^\mu] = 0 \tag{12.3}$$

The first equation is familiar from quantum field theory as the defining equation for the generators of Lorentz group. The third equation is obvious: translations in flat space commute. The second equation isn't immediately obvoius, but may be quickly verified using the matrix form for the generators and then translating the result into index notation.
From now on we omit the $\sim$.

Casimirs (not obvious)
Since the Lie algebra for the Poincare group is not semisimple, the Killing form is not useful (recall that if $L(G)$ is semi-simple, then the Killing form is non-degenerate). However, we do have the following Casimirs (i.e. they commute with both the M and the P operators) in the universal enveloping algebra:

$$P^2 = P_\mu P^\mu$$
$$W^2 = W_\mu W^\mu,$$

where $W_\mu = \epsilon_{\mu\nu\rho\tau} M^{\nu\rho} P^\tau$ ($\epsilon_{0123} = 1$, ϵ totally antisymmetric).
Claim: $W_\mu P^\mu = 0$
Proof: $W_\mu P^\mu = \epsilon_{\mu\nu\rho\tau} M^{\nu\sigma} P^\tau P^\mu$. But $P^\tau P^\mu$ is symmetric in the indices τ, μ while ϵ is totally antisymmetric, thus the contraction vanishes.

12.1 Representations

The Poincare group has finite dimensional representations (e.g., the five-dimensional representation above), but these are not unitary! In quantum mechanics and quantum field theory, particle states transform under symmetries by unitary operators. Thus we need unitary irreps of the Poincare group, which will necessarily be infinite dimensional. We'll find these by working with functions over the homogeneous space $\mathcal{M}_4$.
Recall: A *homogeneous space* for a group G is a non-empty manifold upon with G acts continuously and transitively.
Recall: An action of the group G on X is *transitive* if X is non-empty and for any $x, y \in X$, there exists a $g \in G$ such that ${}^g x = y$. (The left superscripts denotes the action).

12.1.1 Scalar Irreducible Representations

Consider functions of the form $\exp(ikx) = \exp(ik_\mu x^\mu)$. Under the translation $x \longrightarrow x + a$ we have $\exp(ikx) \longrightarrow \exp(ik(x + a)) = \exp(ika)\exp(ikx)$. Thus we get a 1-dimensional irreducible representation of the translation group labelled by k_μ, i.e., by momentum. Under Lorentz transformations $x \longrightarrow Lx$ we have $\exp(ikx) \longrightarrow \exp(ikLx) = \exp(i(kL)x)$. In indices, $k_\mu x^\mu \longrightarrow k_\mu L^\mu{}_\nu x^\nu = (k_\mu L^\mu{}_\nu)x^\nu$. In other words, k undergoes a Lorentz transformation. However, as a Lorentz scalar $k^2 = k_\mu k^\mu$ is conserved. Thus in an *irreducible representation* we can fix $k^2 = m^2$. (Note that k^2 is an eigenvalue of the Casimir P^2.) The following space of functions transforms irreducibly under the Poincare group:

$$\psi(x) = \int f(k)\delta(k^2 - m^2)\exp(ikx)\, d^4k,$$

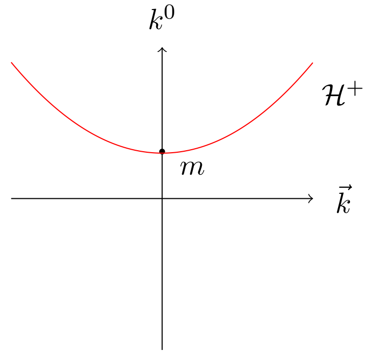
where $f(k)$ is an arbitrary (well-behaved) function. (As we've seen in quantum field theory, the δ-function may be integrated out in integrals like these, leaving an integral over 3-momentum.) f is really a function on $\mathcal{H}^+$, the hyperboloid $k^2 = m^2$ with $k_0 > 0$ in k-space. Thus we shift our focus from $\mathcal{M}_4$ to $\mathcal{H}^+$ instead.

12.1.2 Irreducible Representations with Spin

We can extend the above discussion to allow f to be valued in some vector space over $\mathcal{H}^+$. The key idea is the following: $\mathcal{H}^+$ is a coset space for the Lorentz group:

$$\mathcal{H}^+ = SO(1,3)^\uparrow / SO(3).$$

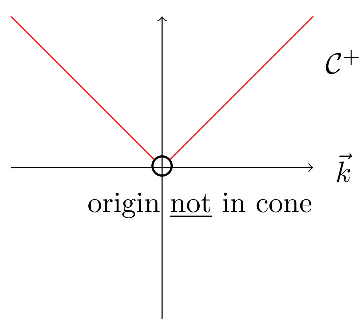
$SO(3)$ is the isotropy group of $k^\mu = (m, \vec{0})$. For an irreducible representation of the Poincare group, we need $SO(3)$ to act irreducibly on the vector-valued function $\vec{f}(m, \vec{0})$. Thus $\vec{f}$ is a state with *definite spin* j (integer or half integer). The basis states $\vec{f}$ are labelled by j and $j_3 = \{-j, -j+1, ..., j-1, j\}$.



Remark. This is an example of a more general construction. Once can construct an infinite dimensional representation of G from functions on G/H valued in a vector space V. There must be an irreducible action d of H on V. This acts as a base point of G/H where H is the isotropy group / stabilizer. The result is called the representation of G *induced* from the irreducible representation d of H. (cf. induced representation on Wikipedia).

12.1.3 Massless Case

If $m = 0$, the orbit of $SO(1,3)^\uparrow$ in k-space is $\mathcal{C}^+$, where $\mathcal{C}^+$ is the cone shown in the picture below, with the origin / point of the cone omitted. One must choose a base point of $\mathcal{C}^+$, e.g., $(|k_0|, 0, 0, k_0)$. Then we see that $\mathcal{C}^+ = SO(1,3)^\uparrow / I$, where I is the non-compact 3-dimensional isotropy subgroup / stabilizer. I has three generators:$K_1 + J_2, K_2 - J_1,$ and J_3.



Theses generators satisfy the following commutation relations:

$$\begin{aligned}
[J_3, K_1 + J_2] &= K_2 - J_1 \\
[J_3, K_2 - J_1] &= -K_1 - J_2 \\
[K_1 + J_2, K_2 - J_1] &= 0.
\end{aligned}$$

All of these follow after one line of computation with the Lorentz generators. Note in particular the final relation, which vanishes. Note: These are the defining relations for the Euclidean group in 2-dimensions, which is usually called $SE(2)$ or $ISO(2)$. This Euclidean group consists of translations in a plane along with rotations about an axis perpendicular to the plane. In our case, I has a compact $SO(2)$ subgroup, the rotations about the axis $(0, 0, k)$. Irreducible representations of I

are labelled by the *helicity* eigenvalue j_3 of $SO(2)$, which is either integer or half-integer. The irreducible representation of I is 1-dimensional:

$$\begin{aligned} d(K_1 + J_2) &= 0 \\ d(K_2 - J_1) &= 0 \\ d(J_3) &= j_3 \end{aligned}$$

Thus a Poincare irreducible representation is constructed from functions $\vec{f}(k)$ on $\mathcal{C}^+$ with helicity j_3. j_3 is the projection of the spin $\vec{J}$ onto the momentum direction. For massless particles, helicity j_3 is a Poincare invariant. There is no meaning for spin in other directions.

- Neutrinos: helicity $\pm\frac{1}{2}$ in the standard model
- Photons: helicity ± 1, i.e., *not* 0. These correspond to circularly polarized photons in opposite directions.

12.2 A Slightly Different Perspective

It is instructive to rephrase these results in slightly different way following the discussion of Weinberg to emphasize the physical importance of some of previous ideas.

12.2.1 One-Particle States

Suppose that we're dealing with a quantum system, so we're interested in the physical state vectors. As we know, the components of the energy-momentum 4-vector all commute with each other, so their eigenvalues form a natural set of numbers by which to characterized the physical states:

$$P^\mu \left|p, \sigma\right\rangle = p^\mu \left|p, \sigma\right\rangle,$$

where the label σ is included to allow for other degrees of freedom in our system. Further, we assume as part of our *definition* of one-particle states that the label σ is discrete.
As usual, we're interested in how the states transform under different unitary transformations, specifically unitary representations of the Poincare or Lorentz group. Let $U(\Lambda, 0) \equiv U(\Lambda)$ be the unitary transformation associated with the homogeneous Lorentz transformation Λ. Evidently $U(\Lambda)$ acts on $|p, \sigma\rangle$ to produce an eigenstate of P with eigenvalue Λp:

$$\begin{aligned} P^\mu U(\Lambda) \left|p, \sigma\right\rangle &= U(\Lambda) \left(U^{-1}(\Lambda) P^\mu U(\Lambda)\right) \left|p, \sigma\right\rangle, \text{ multiplying by 1} \\ &= U(\Lambda) \left({\Lambda^{-1}}_\rho{}^\mu P^\rho\right) \left|p, \sigma\right\rangle \\ &= U(\Lambda) \left(\Lambda^\mu{}_\rho P^\rho\right) \left|p, \sigma\right\rangle, \text{ by defintion of the inverse matrix} \\ &= U(\Lambda) \left(\Lambda^\mu{}_\rho p^\rho\right) \left|p, \sigma\right\rangle, \text{ since we have an eigenstate} \\ &= \Lambda^\mu{}_\rho p^\rho U(\Lambda) \left|p, \sigma\right\rangle \\ &= \Lambda p\, U(\Lambda) \left|p, \sigma\right\rangle, \text{ which was to be be shown} \end{aligned}$$

In particular, we see that $U(\Lambda) |p, \sigma\rangle$ must be a linear combination of the states $|\Lambda p, \sigma'\rangle$:

$$U(\Lambda) \left|p, \sigma\right\rangle = \sum_{\sigma'} C_{\sigma'\sigma}(\Lambda, p) \left|\Lambda p, \sigma'\right\rangle,$$

where $C_{\sigma'\sigma}(\Lambda, p)$ us just the matrix that mixes the states into the relevant superposition and the labels Λ and p indicate that the mixing generally depends on both the 4-momentum p of the original state and the Lorentz transformation Λ. If we express C as a matrix, an irreducible representation will be one in which C is block diagonal. In such an irreducible representation, the invariant subspaces associated with each block will have similar properties, and we can identify them will different types of particles. The question, then, becomes how to get our hands on the irreducible representations.

To this end it useful to consider the sort of functions of p^μ that will be invariant under (proper orthochronous) Lorentz transformations. Clearly, any function of the Lorentz invariant scalar p^2 will itself be invariant. Moreover, for time-like p^μ (i.e., for $p^2 \leq 0$) the sign of p^0 will also be invariant. This amounts to the statement that a Lorentz transformation will not change a positive energy into a non-physical negative energy.

With these invariants in mind, we can choose a "standard" 4-momentum k^μ for each value of p^2 and p^0, and we then write any other 4-momentum of this class as $p^\mu = L^\mu{}_\nu(p)k^\nu$. ($L$ is a "standard Lorentz transformation depending on p). Using the language of the lecture course, this is saying that the Lorentz group acts transitively on the submanifold of momentum space defined by the values of p^2 and p^0.

Using these conventions, we may then *define* the states $|p,\sigma\rangle$ by

$$|p,\sigma\rangle \equiv N(p)U(L(p))\,|k,\sigma\rangle\,.$$

Here $N(p)$ is a normalization factor that is immaterial at the moment.

We now perform an arbitrary (homogeneous) Lorentz transformation $U(\Lambda)$ on the state we defined above:

$$\begin{aligned}
U(\Lambda)\,|p,\sigma\rangle &= N(p)U(\Lambda L(p))\,|k,\sigma\rangle \ \text{, since } U \text{ is a representation}\\
&= N(p)\left[U(L(\Lambda p))U(L^{-1}(\Lambda p))\right]U(\Lambda L(p))\,|k,\sigma\rangle \ \text{, multiplying by “1”}\\
&= N(p)U(L(\Lambda p))\left[U(L^{-1}(\Lambda p)\Lambda L(p))\right]|k,\sigma\rangle
\end{aligned}\tag{12.4}$$

How has this helped us? If we look closely at the square-bracketed term in the last line, we see that it causes the momentum k of our initial state to undergo the following changes: $k \to p \to \Lambda p \to k$. In other words, the bracketed term gives us a subgroup of the (homogeneous) Lorentz group that leaves our standard 4-momentum k invariant. Using the language of the lecture course, this subgroup is the isotropy subgroup / stabilizer of the base point k. This extremely useful subgroup is often referred to as the *Little Group*. Letting W stand for an element of the Little Group, we see immediately that:

$$U(W)\,|k,\sigma\rangle = \sum_{\sigma'} D_{\sigma'\sigma}(W)\,|k,\sigma'\rangle\,,$$

which is just the statement that the Little Group shuffles around the other degrees of freedom to make a superposition of states, all with the original 4-momentum k. One sees immediately that the $D(W)$ is a representation of the Little Group.

Defining $W(\Lambda,p) \equiv L^{-1}(\Lambda p)\Lambda L(p)$, we can combine our results to find:

$$U(\Lambda)\,|p,\sigma\rangle = N(p)\sum_{\sigma'} D_{\sigma'\sigma}(W(\Lambda,p))U(L(\Lambda p))\,|k,\sigma'\rangle$$

We recall our definition $|p,\sigma\rangle \equiv N(p)U(L(p))\,|k,\sigma\rangle$, which allows us to write the previous equation in the form:

$$U(\Lambda)\,|p,\sigma\rangle = \frac{N(p)}{N(\Lambda p)}\sum_{\sigma'} D_{\sigma'\sigma}(W(\Lambda,p))\,\left|\Lambda p,\sigma'\right\rangle$$

Still ignoring the question of normalization for the moment, we see that we've reduced the problem to finding irreducible representations C of the full inhomogeneous Lorentz group to the problem of finding irreducible representations D of the little group W. Since W is a proper subgroup, this new problem should be easier to solve. As the discussion in the lecture mentioned, this representation is known as the representation *induced* by the Little Group.

The next question is what the little groups are for the various standard 4-momenta. The result is given in the table below:

Constraints	k^μ	Little Group W	Interpretation
$p^2 = -M^2 < 0$, $p^0 > 0$	$(M,0,0,0)$	$SO(3)$	Massive Particles
$p^2 = -M^2 < 0$, $p^0 < 0$	$(-M,0,0,0)$	$SO(3)$	Non-physical, negative energy
$p^2 = 0$, $p^0 > 0$	$(\kappa,0,0,\kappa)$	$ISO(2)$	Massless Particles
$p^2 = 0$, $p^0 < 0$	$(-\kappa,0,0,\kappa)$	$ISO(2)$	Non-physical, negative energy
$p^2 = N^2 > 0$	$(0,0,0,N)$	$SO(1,2)$	Non-physical
$p^\mu = 0$	$(0,0,0,0)$	$SO(1,3)$	Vacuum

Of the three physical cases, the vacuum is clearly the most boring: it is invariant under the full Lorentz group. The next easiest to understand is the massive particle. In this case, one picks the standard momentum k^μ to be the one in the rest frame. Clearly only rotations leave a particle at rest, since the boosts by definition change the velocity. Thus the little group is the rotations in three dimensions: $SO(3)$.

The final case of the massless particle requires the most work. However, the arguments that lead to the realization of $ISO(2)$ as the little group are not particularly enlightening from a physical perspective. With this in mind, we omit the details here.